JCMK

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN

EFQM
Recognised for exellence
5 star

Online ISSN 2313-1519
Print ISSN 1812-2892
www.clinmedkaz.org

№22 (6) 2025

Health Care Quality and Patient Satisfaction in Hemodialysis Clinic: Investigation of Patient-Reported Outcomes

> See page 19 and 27

Classification of Combined MLO and CC Mammographic Views Using Vision–Language Models

> See page 89 and 95

Understanding the Healthcare Landscape for Pediatric Epilepsy Patients in Kazakhstan: Challenges and Opportunities

> See page 103 and 108

JCMK

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN

Online ISSN 2313-1519
Print ISSN 1812-2892
№22 (6) 2025
Published since 2004.

EFQM Recognised for exellence 5 star

TÜV CERT Management system as per ISO 9001:2000

EDITORIAL

AIMS AND SCOPE OF THE JOURNAL

Journal “Clinical Medicine of Kazakhstan” (ISSN 1812-2892) is a multi-field dedicated peer-reviewed medical journal. The main thematic scope – publication of materials on medical science and practice, education and healthcare organization. Joint Stock Company “National Scientific Medical Center” publishes the journal bimonthly in a year (in February, April, June, August, October, and December).

All articles sent to editors undergo double-blind review. Manuscripts are judged by two experts exclusively on the basis of their contribution to initial data, ideas and their presentations. Editors accept articles for consideration and publication at no cost. Detailed information is available in the section Information for authors at the end of this material.

The Journal of “Clinical Medicine of Kazakhstan” to the full extent is wedded to initiative of open access and ready to provide free access to full texts of articles, as soon as they will be published and available in the Internet (www.clinmedkaz.org).

Journal was registered in the Ministry of Information of the RK on 05.04.2004 and currently included to the list of Publications, approved by the Committee for Control of Education and Science of the Ministry of Education and Science of the Republic of Kazakhstan for publication of the main outcomes of scientific activity.

The journal is indexed in such international scientific-information bases as Index Copernicus International, Google Scholar, CrossRef. DOAJ.

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN

Scientific and practical journal

JCMK JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN

Editorial Office: Journal of Clinical Medicine of Kazakhstan
Ave Abylay-khan #42, 010009, Astana, Kazakhstan. Phone: +7(7172) 577411
Fax: +7(7172) 232927 E-mail: editor@clinmedkaz.org Web Adress: www.clinmedkaz.org

NATIONAL SCIENTIFIC MEDICAL CENTER JSC, ASTANA CITY,
REPUBLIC OF KAZAKHSTAN

SURGERY AND TRANSPLANTATION

The key activities of these departments of NSMC are reconstructive surgery, transplantation, innovative minimally invasive approaches, microsurgical correction, cell-based therapy of various surgical diseases.
New types of minimally invasive surgical technologies, video-endoscopic and microsurgical techniques, mini-thoracoscopic operations, chemoembolization of tumors were implemented that fundamentally change the approaches to the treatment of many surgical diseases and reduce the time of hospitalization of the patients.

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17360

The Influence of Disease Duration in Pharmacoresistant Focal Epilepsy and Type of Antiepileptic Therapy on Intraoperative Infusion Therapy Parameters

Mukhit Dossov[1], Serik Seitenov[1,2], Nazira Bekenova[3], Galymzhan Kuatbay[4]

[1]Anesthesiology and ICU Department, Medical Center Hospital of the President's Affairs Administration of the Republic of Kazakhstan, Astana Kazakhstan
[2]Deputy head of the medical center, Medical Center Hospital of President's Affairs Administration of the Republic of Kazakhstan, Astana Kazakhstan
[3]Department of Science, Medical Center Hospital of the President's Affairs Administration of the Republic of Kazakhstan, Astana Kazakhstan
[4]Internal Medicine Department, Medical Center Hospital of the President's Affairs Administration of the Republic of Kazakhstan, Astana Kazakhstan

Received: 2025-08-14.
Accepted: 2025-09-25.

J Clin Med Kaz 2024; 22(6): 4-9

Corresponding author:
Mukhit Dossov.
E-mail: dossovmukhit@gmail.com.
ORCID: 0009-0000-3291-362X.

Abstract

Objective: To evaluate the impact of disease duration of pharmacoresistant focal epilepsy and type of antiepileptic therapy on intraoperative infusion therapy parameters.

Materials and Methods: A retrospective cohort study was conducted involving 102 patients (2019–2024) who underwent neurosurgical intervention for pharmacoresistant focal epilepsy. Patients were stratified by disease duration (≤15 years and >15 years) and therapy type (monotherapy and polytherapy). The following parameters were evaluated: infusion volume, infusion therapy composition, fluid balance, urine output, blood loss, extubation time, and length of stay in the intensive care unit.

Results: Polytherapy was associated with significantly higher fluid balance and greater crystalloid volume compared to monotherapy ($p<0.05$). Prolonged disease course (>15 years) was accompanied by a tendency toward increased fluid balance and significantly prolonged time to extubation ($p<0.05$). Hemodynamic parameters remained stable across all groups.

Conclusion: The type of antiepileptic therapy and disease duration are significant factors influencing infusion therapy volume. A personalized approach to infusion management in patients with pharmacoresistant epilepsy may enhance safety and improve outcomes.

Keywords: perioperative fluid management, neurosurgery, epilepsy surgery.

Introduction

Anesthesiologists frequently encounter patients with focal epilepsy undergoing emergency or elective surgical procedures. Infusion therapy represents a crucial component of perioperative management in these patients, as it can directly influence perfusion status, systemic hemodynamics, and the risk of intracranial complications [1].

Patients with long-standing pharmacoresistant focal epilepsy may exhibit persistent autonomic dysregulation (reduced heart rate variability, increased sympathetic tone), which impairs hemodynamic adaptation to stress and increases the risk of perioperative hypotension [1,2]. This necessitates careful selection of infusion regimens.

Antiepileptic drugs may also affect hemodynamics [1,3]. Valproic acid is associated with thrombocytopenia and hyponatremia [4,5], while carbamazepine can precipitate syndrome of inappropriate antidiuretic hormone secretion, thereby increasing the risk of water intoxication [6]. Levetiracetam and lamotrigine less

frequently disrupt fluid and electrolyte balance; however, unpredictable interactions may occur with polytherapy [6,7]. Consequently, such patients require stricter monitoring of fluid balance and electrolytes, as well as early vasopressor administration when signs of autonomic dysfunction are present.

Excessive infusion may lead to cerebral edema and elevated intracranial pressure, while inadequate infusion carries the risk of cerebral hypoperfusion and ischemic injury [8]. Therefore, the anesthesiologist's objective is to provide goal-directed infusion therapy tailored to the individual patient's needs.

Traditional infusion therapy strategies (liberal and restrictive) have inherent limitations, which has led to the development of personalized fluid management concepts. Contemporary perioperative goal-directed therapy (GDT) is based on dynamic variables including stroke volume variation (SVV) and pulse pressure variation (PPV) [9]. However, the implementation of these strategies in clinical practice is limited due to technical requirements (mechanical ventilation, sinus rhythm), inadequate staff training [10], limited data extrapolation [11], and high implementation costs [12]. In settings with limited access to advanced monitoring, clinical practice employs a simplified approach focused on basic markers: mean arterial pressure, central venous pressure, urine output, lactate levels, and acid-base parameters [9,13]. Maintenance of mean arterial pressure is critically important for preventing cerebral hypoperfusion. Vasopressors (norepinephrine, phenylephrine) are utilized when necessary, particularly when hypotension occurs in the context of adequate circulating blood volume [8].

Given the limitations of dynamic monitoring, an individualized fluid therapy protocol based on basic perfusion parameters represents the optimal strategy. This approach allows adaptation to the specific needs of individual patients, particularly in high-risk situations such as neurosurgical interventions for pharmacoresistant epilepsy [14,15].

Thus, the selection of fluid management strategy in patients with pharmacoresistant epilepsy should consider not only standard hemodynamic parameters, but also individual disease characteristics, duration of illness, and type of antiepileptic therapy (AET). A personalized approach to fluid therapy in this patient population becomes not only desirable but a necessary condition for safety and improved treatment outcomes [16]. Based on our analysis of the scientific literature, we found no studies investigating the influence of disease duration and type of antiepileptic therapy on fluid therapy parameters during the intraoperative period. Therefore, the present study aims to evaluate the characteristics of fluid therapy within the framework of personalized fluid management, taking into account epilepsy duration and type of antiepileptic therapy in neurosurgical patients.

Study Objective: To evaluate the influence of disease duration of pharmacoresistant focal epilepsy and type of antiepileptic therapy on fluid therapy parameters during the intraoperative period.

Methods

This study represents a retrospective cohort study conducted at the Medical Centre Hospital of President's Affairs Administration of the Republic of Kazakhstan (MCH PAA RK). The study included 102 patients who underwent surgery for pharmacoresistant focal epilepsy from 2019 to 2024.

Inclusion criteria: age ≥18 years; confirmed diagnosis of pharmacoresistant focal epilepsy; neurosurgical intervention performed with available medical documentation; completed perioperative period (up to 48 hours postoperatively). Exclusion criteria: patients with generalized forms of epilepsy; patients with incomplete data for key parameters; presence of severe comorbid somatic pathology (e.g., end-stage cardiovascular disease, malignancy, end-stage liver or renal disease); bleeding, surgical reintenvention.

The perioperative fluid therapy protocol during surgical interventions for focal epilepsy was based on basic perfusion parameters according to target values: mean arterial pressure (MAP) ≥70 mmHg; central venous pressure (CVP) within 6-12 mmHg; urine output >0.5 mL/kg/h; lactate level <2.0 mmol/L. To achieve and maintain target values, fluid therapy was administered using 0.9% sodium chloride solution; when larger volumes were required and in cases of hyperchloremia, balanced solution Sterofundin (B.Braun Medical, AG, Switzerland) was infused, with Gelofusine (B.Braun Medical, AG, Switzerland) used as the colloidal solution. Within the fluid therapy protocol framework, norepinephrine was administered in the presence of hemodynamic instability (MAP <65 mmHg) despite adequate volume loading.

Recorded parameters: demographic data: age, sex, body mass index (BMI); fluid therapy parameters: total infusion volume, blood loss volume, urine output, calculated fluid balance. Outcome measures: duration of mechanical ventilation, sedation level according to the RASS (Richmond Agitation-Sedation Scale) [17], length of stay in the intensive care unit.

Patients were stratified according to two criteria. Based on disease duration, they were divided into two subgroups: Group 1 — disease duration ≤15 years; Group 2 — >15 years. Additionally, based on the nature of antiepileptic therapy, patients were divided into: monotherapy group (single antiepileptic drug) and polytherapy group (two or more drugs).

Statistical analysis

Statistical processing was performed using SPSS 20 and JASP software. For quantitative variables, means and standard deviations were calculated, with medians and interquartile ranges when appropriate. Categorical data are presented as frequencies and percentages. Group comparisons were performed using analysis of variance or the Kruskal-Wallis test depending on data distribution. P-values <0.05 were considered statistically significant.

Results

The 102 patients included in the study underwent surgery for pharmacoresistant focal epilepsy. The mean age of patients was 33.2±8.2 years. There were 59.8% males and 40.2% females. The mean body mass index was 23.9, and detailed analysis revealed that 7.8% were underweight, 15.7% (16) patients were overweight, 10.8% (11) had class I obesity, and 1 (0.98%) patient had class II obesity.

Patients were stratified by disease duration (≤15 years and >15 years) and by type of antiepileptic therapy (monotherapy and polytherapy). The table presented below reflects differences in clinical and demographic data (Table 1). The age of patients with disease duration >15 years was statistically significantly higher compared to the ≤15 years group (34.1 vs. 30.7, p<0.05). BMI and sex did not differ between subgroups.

The mean age of patients according to therapy type was 32.3 and 32.4 for monotherapy and polytherapy, respectively. Both monotherapy and polytherapy groups were predominantly male. The RASS scale scores after extubation did not differ significantly between groups, with values predominantly close to "-1" (drowsy — sleeps intermittently, awakens to verbal stimulation, responds, then falls asleep again). ICU length of

Table 1 Clinical and demographic data of patients according to therapy type and disease duration

Parameters		Type of antiepileptic therapy		p-value	Disease duration		p-value
		Monotherapy (n=37)	Polytherapy (n=65)		≤15 years (n=53)	>15 years (n=49)	
Mean age (M, 95% CI)		32.3 (29.5-34,9)	32.4 (30.4-34.4)	0.51	30.7 (28.4-33.0)	34.1 (31.9-36.3)	0.01
Body mass index		24.11±4.8	23.68±4.2	0.97	24.92±4.02	23.31±4.69	0.85
Sex	Female	12 (32.4%)	29 (44.6%)		21 (39.6%)	20 (40.8%)	
	Male	25 (67.6%)	36 (55.4%)		32 (60.4%)	29 (59.2%)	
RASS scale score (M±SD)		-0.94±0.22	-0.95±0.21	0.59	-0.96±0.19	-0.93±0.24	0.57
ICU length of stay, hours (M±SD)		14.37±2.8	14.22±3.11	0.71	14.47±3.34	14.06±2.54	0.19
Time to extubation, min (M±SD)		93.45±58.30	97.76±61.35	0.38	81.66±44.16	111.93±70.5	0.03
Duration of anesthesia, min (M±SD)		536.3±116.6	579.59±109.9	0.07	552.29±120.9	559.8±112.3	0.16

M±SD – mean ± standard deviation

Table 2 Fluid therapy parameters according to antiepileptic therapy type and disease duration

Parameters	Disease duration		p-value	Type of antiepileptic therapy		p-value
	≤15 years (n=53)	>15 years (n=49)		Monotherapy (n=37)	Polytherapy (n=65)	
Total infusion volume, mL Me (Q1-Q3)	2800.0 (2400.0-3000.0)	2500.0 (2300.0-3400.0)	0.79	2500 (2200.0-2900.0)	2800.0 (2400.0-3400.0)	0.05
Crystalloids, mL Me (Q1-Q3)	2400.0 (1850.0-2400.0)	2500.0 (2000.0-3000.0)	0.15	2400.0 (1750.0-2500.0)	2500.0 (2000.0-3000.0)	0.04
Colloids, mL Me (Q1-Q3)	500.0 (500.0-625.0)	500.0 (500.0-500.0)	0.78	500.0 (500.0-625.0)	500.0 (500.0-500.0)	0.92
Balanced solutions, mL Me (Q1-Q3)	500.0 (500.0-500.0)	500.0 (500.0-1000.0)	0.35	500.0 (500.0-625.0)	500.0 (500.0-500.0)	0.6
Urine output, mL Me (Q1-Q3)	1000.0 (650.0-1650.0)	900.0 (525.0-1250.0)	0.27	900.0 (550.0-1500.0)	1425.0 (100.0-2040.0)	0.89
Blood loss, mL Me (Q1-Q3)	100.0 (100.0-200.0)	100.0 (100.0-200.0)	0.16	100.0 (100.0-200.0)	100.0 (100.0-200.0)	0.72
Fluid balance, mL Me (Q1-Q3)	1400.0 (960.0-1825.0)	1600.0 (1100.0-2025.0)	0.11	1250.0 (950.0-1650.0)	1700.0 (1150.0-2000.0)	0.02
Norepinephrine dose, μg/kg/min Me (Q1-Q3)	0.040 (0.040-0.050)	0.050 (0.030-0.050)	0.6	0.050 (0.040-0.050)	0.040 (0.040-0.050)	0.69

Me – median. (Q1-Q3) – interquartile range. p – Mann-Whitney test.

stay was >12 hours in all groups with no significant differences detected. Duration of mechanical ventilation after surgery did not differ between antiepileptic therapy type groups, but among disease duration groups, patients with longer disease history required longer postoperative mechanical ventilation ($p<0.05$). Duration of anesthesia did not differ between disease duration groups or antiepileptic therapy type groups.

The results of fluid therapy parameter analysis are presented in the following table (Table 2).

Our study results demonstrated statistically significant differences in net fluid balance between patients according to therapy type. Patients receiving multiple medications had higher fluid balance than patients on monotherapy. In the disease duration group, fluid balance was higher in patients with disease duration <15 years, but this difference was not statistically significant. When analyzing the structure of fluid therapy, the volume of crystalloid solutions was also higher in patients with polytherapy. The volumes of balanced and colloidal solutions administered were approximately equal and showed no significant differences. Regarding blood loss, there were no statistically significant differences in blood loss volumes between polytherapy and monotherapy patients. None of the patients included in the study received blood transfusions. Although urine output volume was higher in polytherapy patients compared to monotherapy patients, no statistically significant differences in urine output volume between groups were identified.

In monotherapy patients, the median fluid balance was approximately 1250 mL, with values ranging from ~950 to 1650 mL. In polytherapy patients, the median was higher (approximately 1700 mL). Patients receiving polytherapy were characterized by higher intraoperative infusion volumes and fluid balance with wider value distribution.

Hemodynamic parameters (systolic blood pressure, diastolic blood pressure, mean arterial pressure, heart rate, SpO_2) remained stable throughout all surgical stages, regardless of disease duration or therapy type. Mean values were within physiological normal ranges. No statistically significant differences between groups were identified.

To assess the association between antiepileptic therapy type (mono- or polytherapy), disease duration, and clinical parameters, odds ratios (OR) with 95% confidence intervals (CI) were calculated, and χ^2 analysis with corresponding p-values was performed (Table 3).

Table 3 Association of outcome parameters with antiepileptic therapy type and disease duration

Parameters	Therapy Type			Disease Duration		
	OR (95% CI)	χ^2	p	OR (95% CI)	χ^2	p
RASS scale	1.18 (0.19-7.41)	0.03	0.86	0.60 (0.10-3.76)	0.30	0.58
ICU length of stay	0.66 (0.24-1.78)	0.69	0.41	0.58 (0.23-1.57)	1.33	0.25
Time to extubation	1.18 (0.52-2.68)	0.16	0.69	1.86 (0.84-4.12)	2.39	0.12
Intraoperative fluid balance	0.48 (0.29-0.80)	5.43	0.2	0.32 (0.08-1.23)	2.89	0.09

OR – Odds Ratio. χ^2 – Pearson's chi-square test.

Sedation level (RASS scale) after extubation was not associated with either therapy type or disease duration. Wide confidence intervals indicate high uncertainty in these estimates.

When analyzing ICU length of stay, odds ratios below 1.0 suggest a tendency toward shorter stays with monotherapy; however, the differences do not reach statistical significance.

A trend toward increased extubation time was noted in patients with longer epilepsy history, though the differences were not statistically significant. The only parameter demonstrating a statistically significant association was therapy type and net fluid balance: patients on polytherapy had higher net fluid balance volumes. The association with disease duration also showed a trend but did not reach statistical significance. Our study results demonstrated that monotherapy is associated with lower net fluid balance requirements compared to polytherapy.

Discussion

This study demonstrates the characteristics of fluid loading in patients undergoing surgery for pharmacoresistant focal epilepsy, considering disease duration and type of antiepileptic therapy (AET).AET type emerged as an independent predictor of increased net fluid balance. Patients on polytherapy received statistically significantly higher infusion volumes, including crystalloids, compared to monotherapy patients. In both subgroups (by disease duration), patients receiving polytherapy showed a trend toward greater intraoperative fluid therapy volumes. This may indicate increased fluid requirements, differences in vascular status, or variations in anesthetic management strategies. In the study by Sae-Phua V. et al. [18], craniotomy patients received larger intraoperative volumes without goal-directed fluid therapy, predominantly crystalloids, but this study was conducted in elderly patients. Mishra N. and colleagues [19] optimized fluid balance through perioperative goal-directed fluid therapy with dynamic hemodynamic parameter monitoring. Studies by Luo J. et al. [16] emphasize the importance of goal-directed therapy, which enabled them to reduce fluid loading to 2.8 mL/kg/hour. Sundaram C.S. et al. [20], studying fluid loading in patients undergoing brain tumor surgery, noted that CVP-guided management resulted in higher volume infusion. In our study, maximum infusion volumes were similar when CVP-guided approaches were used. The fluid therapy structure across all studies consisted of crystalloids and colloids, with vasoactive agents applied for hemodynamic stabilization. In our study, norepinephrine was similarly used for hemodynamic stabilization when fluid therapy proved ineffective and did not exceed average therapeutic doses. These data are consistent with results from several studies [21,22], which emphasize the necessity of a balanced approach to vasoactive therapy in neurosurgical operations.

Our data confirm that prolonged epilepsy history (>15 years) is associated with increased net fluid balance and prolonged time to extubation. This may reflect more pronounced systemic and metabolic disturbances, requiring more cautious anesthetic management of such patients. These findings are comparable to data from C. Rajkalyan [23] and Hrdy et al. [24], who also emphasize the necessity of stricter fluid volume monitoring during neurosurgical interventions. Mishra et al. [19] report that individualized fluid therapy reduces mechanical ventilation time and hospitalization duration, which aligns with our conclusions regarding the relationship between fluid loading and duration of ventilatory support [25].

Increased extubation time was identified in patients with prolonged epilepsy disease. These data coincide with observations by Bindra A. [25] and Bloor M. [26], where patients with extended disease history more frequently required prolonged mechanical ventilation. Larkin et al. [27] described that combination regimens are more commonly used in resistant patients and recommend more thorough perioperative assessment. The significant association of increased net fluid balance with polytherapy may be related to the combined effects of medications on the central nervous system and vascular tone regulation [26].

According to our data, awakening duration and sedation level assessment using the RASS scale showed no association with therapy type or disease duration. Ouchi K. [28] notes that combination antiepileptic therapy prolongs anesthetic effects and consequently awakening time. Works by numerous authors describe significant complexities in anesthetic and perioperative management of epilepsy patients, including the awakening phase and seizure provocation prevention [25–30].

These data emphasize the necessity of an individualized approach to fluid therapy and hemodynamic control during neurosurgical interventions in epilepsy patients, particularly in the presence of polytherapy and prolonged disease history. Our findings confirm the importance of a comprehensive approach considering therapy characteristics and disease duration when planning fluid management strategies and perioperative care. This supports the need for personalized approaches and potential integration of goal-directed fluid therapy protocols into clinical practice, especially in patients at high risk for metabolic disturbances and hemodynamic instability.

The study has limitations inherent to its retrospective design, limited sample size, and absence of invasive hemodynamic and tissue perfusion monitoring. Additionally, antiepileptic drug doses and pharmacokinetics were not considered, nor were specific surgical nuances. However, the obtained data provide a foundation for further prospective studies incorporating goal-directed fluid therapy with lactate monitoring and dynamic hemodynamic parameter monitoring, as well as analysis of long-term neurological outcomes.

Conclusion

The obtained results demonstrate that patients with prolonged disease history and polytherapy are characterized by

higher fluid loading and positive net fluid balance. Antiepileptic therapy type and disease duration are independent predictors of fluid loading. These parameters may indicate more pronounced metabolic disturbances and the necessity for fluid management strategy adaptation in this subgroup.

Author Contributions: Conceptualization, D.M.; methodology, D.M. and S.S.; validation, D.M. and B.N.; formal analysis, N.B., G.K. and D.M.; investigation, D.M. and N.B.; resources, S.S.; data curation, S.S. and D.M..; writing – original draft preparation, D.M. and G.K.; writing – review and editing, D.M.,B.N. and G.K.; visualization, N.B. and G.K.; supervision, S.S.; project administration, D.M.; funding acquisition, D.M. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. Perks A, Cheema S, Mohanraj R. Anaesthesia and epilepsy. *Br J Anaesth.* 2012;108:562-571. https://doi.org/10.1093/bja/aes027
2. Cascella M, Bimonte S, Di Napoli R. Delayed Emergence from Anesthesia: What We Know and How We Act. *Local Reg Anesth*. 2020 Nov 5;13:195-206. https://doi.org/10.2147/LRA.S230728
3. Scott MJ, APSF Hemodynamic Instability Writing Group. Perioperative Patients With Hemodynamic Instability: Consensus Recommendations of the Anesthesia Patient Safety Foundation. *Anesth Analg.* 2024;138:713-724. https://doi.org/10.1213/ANE.0000000000006789
4. Jaber M, Shawahna R, Abu-Issa M, Radwan F, Dweik M. Anesthesia considerations for patients with epilepsy: Findings of a qualitative study in the Palestinian practice. *Epilepsy Behav.* 2021;123:108278. https://doi.org/10.1016/j.yebeh.2021.108278
5. Kodama M, Higuchi H, Ishii-Maruhama M, Sizuki E, Miyao H. Multi-drug therapy for epilepsy influenced bispectral index after a bolus propofol administration without affecting propofol's pharmacokinetics: a prospective cohort study. *Sci Rep.* 2020;10:1578. https://doi.org/10.1038/s41598-020-58460-2
6. Larkin CM, O'Brien DF, Maheshwari D. Anaesthesia for epilepsy surgery. *BJA Educ.* 2019;19:383-389. https://doi.org/10.1016/j.bjae.2019.08.001
7. Hamdi H, Kishk N, Shamloul R, Moawad MK, Baghdadi M, Rizkallah M, Nawito A, Mohammad ME, Nazmi H, Nasr YM, Waly SH, Elshahat M, Magdy R, Othman AS, Nafea H, Fouad AM, Elantably I, Rizk H, Elsayyad E, Morsy AA. Resective epilepsy surgery in a limited-resource settings: A cohort from a multi-disciplinary epilepsy team in a developing country. *Surg Neurol Int.* 2023;14:240. https://doi.org/10.25259/SNI_1081_2022
8. Kutum C, Lakhe P, Ghimire N, Bc AK, Begum U, Singh K. Intraoperative goal-directed fluid therapy in neurosurgical patients: A systematic review. *Surg Neurol Int.* 2024;15:233. https://doi.org/10.25259/SNI_412_2024
9. Michard Frederic et al. Goal-directed haemodynamic therapy: what else? *British Journal of Anaesthesia.* 2022;128:416-433. https://doi.org/10.1016/j.bja.2022.01.011
10. MacDonald N, Pearse RM. Are we close to the ideal intravenous fluid? *Br J Anaesth.* 2017 Dec 1;119(suppl_1):i63-i71. https://doi.org/10.1093/bja/aex293. Solares GJ, Garcia D, Monge Garcia MI, Gracia E, Corrales G, Lopez J, Granell M, Gil A, Gutierrez P, Gayoso MS, Torrado H, Salgado P, Suarez-Sipmann F, Gomar C, Pelosi P, Schultz MJ, Tusman G. Real-world outcomes of the hypotension prediction index in the management of intraoperative hypotension during non-cardiac surgery: a retrospective clinical study. *J Clin Monit Comput.* 2023;37:211-220. https://doi.org/10.1007/s10877-022-00881-7
11. Benes J, Zatloukal J, Simanova A, Chytra I, Kasal E, Svitak R, Pradl R, Stepan M. Cost analysis of the stroke volume variation guided perioperative hemodynamic optimization – an economic evaluation of the SVVOPT trial results. *BMC Anesthesiol.* 2014;14:40. https://doi.org/10.1186/1471-2253-14-40
12. Futier E, Lefrant JY, Guinot PG, Godet T, Lorne E, Cuvillon P, Bertran S, Leone M, Pastene B, Piriou V, Molliex S, Albanese J, Julia JM, Tavernier B, Imhoff E, Bazin JE, Constantin JM, Pereira B, Jaber S, INPRESS Study Group. Effect of Individualized vs Standard Blood Pressure Management Strategies on Postoperative Organ Dysfunction Among High-Risk Patients Undergoing Major Surgery: A Randomized Clinical Trial. *JAMA*. 2017;318:1346-1357. https://doi.org/10.1001/jama.2017.14172
13. Kendrick JB, Kaye AD, Tong Y, Belani K, Urman RD, Hoffman C, Liu H. Goal-directed fluid therapy in the perioperative setting. *J Anaesthesiol Clin Pharmacol.* 2019;35:S29-S34. https://doi.org/10.4103/joacp.JOACP_26_18
14. Kaufmann T, Clement RP, Scheeren TWL, Saugel B, Keus F, van der Horst ICC. Perioperative goal-directed therapy: A systematic review without meta-analysis. *Acta Anaesthesiol Scand.* 2018 Nov;62(10):1340-1355. https://doi.org/10.1111/aas.13212. Epub 2018 Jul 6. PMID: 29978454.
15. Luo J, Xue J, Liu J, Liu B, Liu L, Chen G. Goal-directed fluid restriction during brain surgery: a prospective randomized controlled trial. *Ann Intensive Care.* 2017;7:16. https://doi.org/10.1186/s13613-017-0239-8
16. Sessler CN, Gosnell MS, Grap MJ, Brophy GM, O'Neal PV, Keane KA, Tesoro EP, Elswick RK Jr. The Richmond Agitation–Sedation Scale. *Am J Respir Crit Care Med.* 2002;166:1338-1344. https://doi.org/10.1164/rccm.2107138
17. Sae-Phua V, Tanasittiboon S, Sangtongjaraskul S. The Effect of Goal-directed Fluid Management based on Stroke Volume Variation on ICU Length of Stay in Elderly Patients Undergoing Elective Craniotomy: A Randomized Controlled Trial. *Indian J Crit Care Med.* 2023;27:709-716. https://doi.org/10.5005/jp-journals-10071-24551

18. Mishra N, Rath GP, Bithal PK, Chaturvedi A, Chandra PS, Borkar SA. Effect of Goal-Directed Intraoperative Fluid Therapy on Duration of Hospital Stay and Postoperative Complications in Patients Undergoing Excision of Large Supratentorial Tumors. *Neurol India*. 2022 Jan-Feb;70(1):108-114. https://doi.org/10.4103/0028-3886.336329 PMID: 35263862.
19. Sundaram SC, Salins SR, Kumar AN, Korula G. Intra-operative fluid management in adult neurosurgical patients undergoing intracranial tumour surgery: randomised control trial comparing pulse pressure variance (PPV) and central venous pressure (CVP). *J Clin Diagn Res.* 2016;10:UC01-5. https://doi.org/10.7860/JCDR/2016/18377.7850
20. Ryu T. Fluid management in patients undergoing neurosurgery. *Anesth Pain Med (Seoul)*. 2021;16:215-224. https://doi.org/10.17085/apm.21072
21. Hasanin A, Zanata T, Osman S, Abdelwahab Y, Samer R, Mahmoud M, Elsherbiny M, Elshafaei K, Morsy F, Omran A. Pulse Pressure Variation-Guided Fluid Therapy during Supratentorial Brain Tumour Excision: A Randomized Controlled Trial. *Open Access Maced J Med Sci.* 2019;7:2474-2479. https://doi.org/10.3889/oamjms.2019.682
22. Rajkalyan C, Tewari A, Rao S, Avitsian R. Anesthetic considerations for stereotactic electroencephalography implantation. *J Anaesthesiol Clin Pharmacol*. 2019;35:434-440. https://doi.org/10.4103/joacp.JOACP_342_18
23. Hrdy O, Duba M, Dolezelova A, Roskova I, Hlavaty M, Traj R, Bönisch V, Smrcka M, Gal R. Effects of goal-directed fluid management guided by a non-invasive device on the incidence of postoperative complications in neurosurgery: a pilot and feasibility randomized controlled trial. *Perioper Med (Lond).* 2023;12:32. https://doi.org/10.1186/s13741-023-00321-3
24. Bindra A, Chouhan RS, Prabhakar H, Bithal PK, Dash HH, Chandra PS. Perioperative anesthetic implications of epilepsy surgery: a retrospective analysis. *J Anesth.* 2015;29:229-234. https://doi.org/10.1007/s00540-014-1919-2
25. Bloor M, Nandi R, Thomas M. Antiepileptic drugs and anesthesia. *Paediatr Anaesth.* 2017;27:248-250. https://doi.org/10.1111/pan.13074
26. Ouchi K. The number and kind of antiepileptics affect propofol dose requirement for anesthesia: observational study. *Odontology*. 2020;108:102-108. https://doi.org/10.1007/s10266-019-00457-z
27. Bhosale UA, Yegnanarayan R, Gupta A, Shah P, Sardesai S. Comparative pre-emptive analgesic efficacy study of novel antiepileptic agents gabapentin, lamotrigine and topiramate in patients undergoing major surgeries at a tertiary care hospital: a randomized double blind clinical trial. *J Basic Clin Physiol Pharmacol.* 2017;28:59-66. https://doi.org/10.1515/jbcpp-2016-0013
28. Li W, Lai W, Peng A, Deng K, Fang H, Ren C, Deng Y. Two cases of anesthetics-induced epileptic seizures: a case report and literature review. *Acta Epileptologica.* 2022;4:5. https://doi.org/10.1186/s42494-021-00074-w

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17357

The Silent Threats of Aging: Fear of Fragile Bones, Stiff Arteries, or Time's Inevitable Betrayal? An Age-Matched Study

Maja Avramovska[1,2], Petar Avramovski[3], Liljana Todorovska[4], Biljana Taleva[5], Kosta Sotiroski[6], Dejan Zdraveski[7]

[1]Department of Obstetrics and Gynecology Clinical Hospital D-r Trifun Panovski, North Macedonia
[2]Department of Higher Medical School, University St. Clement of Ohrid, Bitola, North Macedonia
[3]Department of Internal Medicine, Clinical Hospital D-r Trifun Panovski, Bitola, North Macedonia
[4]Department of Nuclear Medicine, Clinical Hospital "D-r Trifun Panovski" – Bitola, Bitola, North Macedonia
[5]Department of Children Surgery, "Mother Teresa Clinical Center – Skopje", Skopje, North Macedonia
[6]Depatment of Statistics, University St. Clement of Ohrid, Bitola, Faculty of Economics – Prilep, North Macedonia.
[7]Faculty of economics, University St. Clement of Ohrid, Bitola, Prilep, North Macedonia

Received: 2025-09-12.
Accepted: 2025-10-18.

J Clin Med Kaz 2025; 22(6): 10–18

Corresponding author:
Petar Avramovski
E-mail: avramovski@gmail.com.
ORCID: 0000-0003-2816-3365.

Abstract

Introduction: This study assessed the predictive power of femoral neck (FN) bone mineral density (BMD) and pulse wave velocity (PWV) for fracture and cardiovascular-related mortality over a three-year follow-up in a representative cohort.

Methods: A total of 142 participants (54 males, 38%), aged 56 ± 7.2 years, were enrolled in this prospective observational study. FN BMD was measured using dual-energy X-ray absorptiometry (DXA), and carotid-to-femoral PWV was determined via Doppler ultrasound.

Results: Mean PWV was significantly higher in non-survivors compared to survivors (10.9 ± 3.2 m/s vs. 8.6 ± 2.1 m/s, p = 0.0041). FN BMD was lower in non-survivors (0.658 ± 0.131 g/cm²) than in survivors (0.852 ± 0.150 g/cm², p = 0.002). Logistic regression identified PWV as a strong determinant of mortality [coefficient: 0.1593; odds ratio (OR): 1.17; 95% confidence interval (CI): 1.04–1.32; p = 0.01], while FN BMD also showed significance (coefficient: –6.6336; OR: 0.0013; 95% CI: 0.000–0.156; p = 0.0064). However, in age-matched analysis, only PWV remained significant (OR: 2.77; 95% CI: 1.70–4.51; p < 0.0001]. Receiver operating characteristic (ROC) analysis demonstrated superior predictive accuracy for PWV [area under the curve (AUC): 0.958; cutoff: 11.3 m/s; sensitivity: 94.6%; specificity: 88.1%] compared with FN BMD (AUC: 0.560).

Conclusion: PWV showed outstanding accuracy for predicting all-cause mortality, outperforming FN BMD and remaining independent of age. These findings establish PWV as a robust prognostic marker for mortality risk, highlighting its potential role in improving clinical risk stratification for vascular aging and cardiovascular outcomes.

Keywords: Femoral Neck; Pulse Wave Velocity; Osteoporosis; Bone health; Arterial Stiffness.

Introduction

Aging is a complex and inevitable process that brings about various physiological changes, many of which occur silently, posing significant risks to health and quality of life [1]. Among these silent threats, two stand out for their profound impact: osteoporosis, characterized by gradual bone deterioration and a heightened risk of fractures, and arterial stiffness, a key factor in cardiovascular complications and increased mortality [2]. Both conditions are prevalent in aging populations and share common risk factors, including chronic inflammation, hormonal changes, renal function, nutritional deficiencies, genetic predispositions, and physical activity levels, and

diabetes [3]. Despite their different manifestations, osteoporosis and arterial stiffness may act independently or synergistically, increasing the likelihood of adverse outcomes such as fractures and cardiovascular death [2, 4]. Calcium mobilization from weakened bones to stiffer arteries highlights the interplay between osteoporosis and arterial stiffness in aging [5]. This process, involving calcium transfer to arterial walls, exacerbates bone fragility and promotes vascular calcification, linking these conditions through shared pathophysiological pathways [5, 6].

Or should we be more concerned about rigid and stiffer arteries, which silently escalate the risk of cardiovascular events and sudden death [6]? Alternatively, is it the natural aging process itself—the inevitable decline that encompasses both skeletal and vascular health—that poses the greatest threat? Each of these factors contributes to a broader picture of vulnerability, but understanding their individual and combined effects is essential for developing targeted strategies for risk reduction and treatment. The ramification of progressive arterial stiffening on cardiovascular outcomes is comprehensively established, but its influence on other health aspects in older adults, including bone and muscle conditions, remains less explored [6]. While arterial stiffening, bone mass decline, and muscle wasting are all age-associated changes influenced by overlapping risk factors, it remains uncertain whether they occur as independent parallel processes or stem from shared underlying mechanisms [6, 7]. Osteoporosis, often called the "silent disease," advances quietly without noticeable symptoms until a fracture occurs, resulting in significant health complications and a marked decline in quality of life [8].

Likewise, arterial stiffness, commonly assessed through pulse wave velocity (PWV), serves as a pivotal determinant of cardiovascular hazard and survival likelihood, frequently advancing silently until severe complications develop [9]. Early identification and management of these risks can greatly enhance health outcomes and lessen the impact of aging-related conditions. Bone mineral density (BMD) gradually decreases with age, mainly due to an imbalance in bone turnover.

In postmenopausal women, this decline speeds up as a result of lower estrogen levels, heightening the likelihood of developing osteopenia and osteoporosis [10]. In aging individuals, reduced BMD compromises bone integrity, increasing fracture susceptibility, particularly in the vertebrae, femur, and wrist [10, 11]. In the general population, arterial stiffening occurs gradually with age, driven by factors such as declining vascular elasticity and cumulative exposure to cardiovascular risk factors [12].

This manuscript aims to present findings from a three-year follow-up study conducted in a general population cohort, focusing on the long-term outcomes associated with osteoporosis and arterial stiffness. Specifically, the study examines the incidence of fractures and cardiovascular death, evaluating the predictive power of BMD and PWV. By analyzing these data, this research seeks to provide valuable insights into the interplay between skeletal and vascular health, emphasizing the importance of early detection and intervention in mitigating the risks associated with aging. In this study, a particular focus will be placed on distinguishing age as an inevitable process from the quality of bone and vascular health in relation to lethal outcomes during a three-year observational period, with the cohort being age-matched. This study highlights the need for a comprehensive approach to managing age-related conditions, addressing both bone and cardiovascular health to improve patient outcomes.

Methods

This prospective observational study, carried out at the Regional Hospital in Bitola, aimed to investigate the relationship between bone health and vascular rigidity in a general population. A total of 142 participants, comprising 54 males (38%), underwent assessments of bone mineral density using dual-energy X-ray absorptiometry (DXA) and arterial stiffness using carotid-to-femoral pulse wave velocity (PWV). The median age and body mass index (BMI) in the studied cohort were 56 years and 27.34 kg/m², in that order. Common comorbidities included hypertension (27.46%), smoking (38.73%), and diabetes (17.6%). The study monitored clinical outcomes over a three-year observation period, spanning from February 2, 2021, to March 15, 2024.

Table 1 Demographic Characteristics, Clinical Biomarkers, and Outcomes Comparison in Cardiovascular Death and Femoral Neck Fracture Fatal Outcomes Based on bone mineral density (BMD and pulse wave velocity (PWV) Values

Variables	All	CARDIOVASCULAR (PWV)		P	F_NECK fatal outcome (BMD)		P
	N = 142	Survived	Non-survived		Survived	Non - survived	
		N = 134	N = 8		N = 136	N = 6	
Age, years	56 (49 - 68)	57.47 ± 11.35	69.5 ± 6.43	0.0036	57.15 ± 11.26	72.67 ± 3.44	0.001
BMI, kg/m2	27.34 (24.15 to 31.18)	27.79 ± 4.32	27.32 ± 3.95	0.764	27.87 ± 4.39	25.29 ± 2.41	0.156
Hypertension, N (%)	39 (27.46%)	36 (26.47)	3 (37.5)	0.462	37 (27.21)	2 (33.330	0.744
Smoking, N (%)	55 (38.73)	52 (38.8)	3 (37.5)	0.005	53 (38.9)	2 (33.33)	0.075
Diabetes, N (%)	25 (17.6)	21 (15.7)	4 (50)	0.0137	26 (19.11)	3 (50)	0.0672
Fneck fracture (fatal outcome), N (%)	6 (4.22)	/		136	6	0.584	
Cardiovascular death endpoint, N (%)	8 (5.63)	134	8		/		
Fneck survival, months	35.12 ± 4.51	/	0.0035	36	32.3 ± 8.98	0.0013	
Cardiovascular survival, months	34.84 ± 5.42	36	34.94 ± 5.22		/		
BMD Fneck, g/cm2	0.837 (0.733 to 0.927)	0.846 ± 0.157	0.803 ± 0.082	0.803	0.852 ± 0.150	0.658 ± 0.131	0.002
PWV, m/s	9.4 (8.4 to 10.8)	8.6 ± 2.1	10.9 ± 3.2	0.0041	10.49 ± 2.7	11.1 ± 2.7	0.589

The results are presented as: mean and ± standard deviation (SD), median and 25th to 75th percentiles, number (N) and percent (%). BMI, body mass index; Fneck, femoral neck; BMD, bone mineral density and PWV, pulse wave velocity.

Notably, six fatal FN fractures (4.22%) and eight fatal cardiovascular events (5.63%) were documented, including three cases of acute myocardial infarction, two of ischemic stroke, two of sudden cardiac death, and one of decompensated heart failure. Fatal cardiovascular events were defined as deaths directly attributable to cardiovascular causes, confirmed by hospital records, death certificates, or autopsy reports where available. The six fatal FN fracture cases were due to non-cardiovascular complications: one patient died from bronchopneumonia due to prolonged immobilization, two patients succumbed to septicemia resulting from postoperative wound infections, one patient experienced acute respiratory failure, one developed sepsis following a urinary tract infection, and one suffered acute kidney injury with multi-organ failure.

The mean survival period for FN fracture endpoints was 35.12 months, while it was 34.84 months for cardiovascular-related events. The remaining patient characteristics and medical parameters are presented in Table 1, along with their comparison between survived and non-survived patients in both groups (with cardiovascular death and FN fracture death).

Bone mineral density of the FN (FN BMD) was measured using DXA, while carotid-to-femoral PWV propagation was evaluated using Doppler techniques. Carotid-to-femoral PWV was measured using color Doppler ultrasound with electrocardiogram (ECG) synchronization in a quiet, temperature-controlled room after a 10-minute resting period, with participants in the supine position and in a fasting state. The pulse transit time was determined as the time delay between the foot of the Doppler waveform at the carotid and femoral arteries, referenced to the R-wave of the ECG. The distance was measured as the direct surface distance between the carotid and femoral recording sites. Two consecutive measurements were obtained, and the mean value was used for analysis. The methods of DXA and carotid-to-femoral PWV measurement, as described in previous studies [2, 10–15], include determining arterial stiffness by evaluating the duration needed for the arterial pulse signal to travel across the carotid and femoral arteries (cfPWV), dividing the pulse transit time by the known distance between these two points. Patients with chronic unregulated diabetes on insulin therapy, rheumatoid arthritis, pulmonary disorders, malignancies, liver disease, or other persistent diseases affecting bone or heart health—such as recent and past cardiovascular events, heart attacks, or peripheral artery disease—were excluded from the study.

Statistical analysis

The dataset was processed using MedCalc® Statistical Software version 22.002 (MedCalc Software Ltd, Ostend, Belgium, https://www.medcalc.org; 2023). To compare continuous variables, an independent t-test was used for normally distributed data, whereas a non-parametric Mann–Whitney U test was applied for skewed distributions. We used age-matched residual-based matching to estimate the risk for cardiovascular fatal outcomes and FN fracture fatal outcomes. Age adjustment was performed using a residual-based matching approach, where the residuals from a linear regression model of each variable on age were computed and used in subsequent analyses, effectively removing the linear effect of age. Survival probabilities and the number at risk for cardiovascular fatal outcomes and FN fracture fatal outcomes were estimated across censored event-time analysis using the Kaplan-Meier method. To assess the model's capacity to differentiate surviving and non-surviving individuals, diagnostic performance was assessed using ROC (receiver operating characteristic) curves, identifying sensitivity, specificity, and optimal cutoff values. A logistic regression analysis was undertaken to determinate the predictive significance of PWV, FN BMD, and age concerning all-cause mortality.

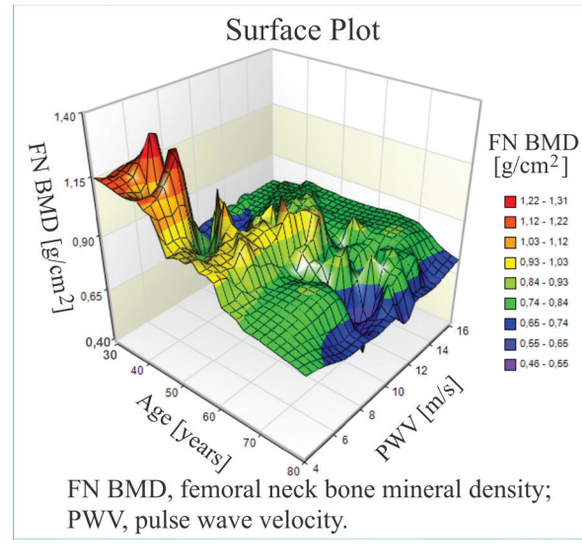
Results

The study evaluated 142 general population participants, examining relationships between femoral neck bone mineral density (FN BMD), pulse wave velocity (PWV), and clinical outcomes over a 36-month follow-up. Demographic, clinical, and procedural parameters, alongside survival outcomes, are summarized in Table 1, comparing survivors and non-survivors for both cardiovascular and FN fracture outcomes.

Patients who died from cardiovascular events were significantly older (69.5 ± 6.43 years, P = 0.0036), and those with fatal FN fractures were also older (72.67 ± 3.44 years, P = 0.001). BMI did not differ significantly for cardiovascular (P = 0.764) or FN fracture outcomes (P = 0.156). Hypertension prevalence was 27.46%, without significant group differences (P = 0.462 for cardiovascular, P = 0.744 for FN fractures). Smoking (38.73%) was significantly associated with cardiovascular mortality (P = 0.005) but not FN fractures (P = 0.075). Diabetes (17.6%) correlated with cardiovascular deaths (P = 0.0137) but not FN fracture mortality.

During follow-up, eight patients (5.63%) died from cardiovascular causes and six (4.22%) from FN fractures. Mean survival was 34.94 ± 5.22 months for cardiovascular deaths and 32.3 ± 8.98 months for FN fracture deaths (P = 0.0027). FN fracture patients had lower FN BMD (0.658 ± 0.131 g/cm² vs. 0.852 ± 0.150 g/cm², P = 0.002). Cardiovascular deaths had higher PWV (10.9 ± 3.2 m/s vs. 8.6 ± 2.1), and FN fracture patients also showed elevated PWV (11.1 ± 2.7 vs. 10.49 ± 2.7). These results indicate strong associations of age, arterial stiffness, and bone health with cardiovascular and FN fracture risk.

Figure 1 shows a 3D surface plot of FN BMD, PWV, and age, illustrating BMD decline with age and higher PWV, emphasizing interactions among vascular rigidity, aging, and bone health. Pearson correlations revealed a strong positive relationship between PWV and age (r = 0.765, p < 0.0001) and moderate negative correlations between PWV and FN BMD



FN BMD, femoral neck bone mineral density; PWV, pulse wave velocity.

Figure 1 – 3D Surface plot depicting the relationship between bone mineral density of the Femoral Neck, pulse wave velocity, and age in the study population

(r = −0.414, p < 0.0001) and age and FN BMD (r = −0.414, p < 0.0001).

Table 2 presents logistic regression results for all-cause mortality. Age was significant [P = 0.008, odds ratio (OR) = 1.071] but partially confounded by PWV due to multicollinearity (variance inflation factor, VIF = 5.732). Lower FN BMD was associated with higher mortality risk (P = 0.0064, OR = 0.0013, b = -6.6336), with a wide confidence interval (0.0000–0.1555).

Table 2 Predictive Value of Pulse Wave Velocity, Femoral Neck Bone Mineral Density and age for All-Cause Mortality

Backward Logistic Regression						
Number of events	14	9.86%	8 (cardiovascular events) + 6 (Femur neck fatal outcome)			
Number censored	128	90.14%				
Total number of cases	142	100.00%				
Null model -2 Log Likelihood	137.858					
Full model -2 Log Likelihood	128.839					
Chi-squared	9.02					
DF	1					
Significance level	P = 0.0027					
Covariate	Coefficient	SE	Wald	P	OR	95% CI of OR
PWV (m/s)	0.1593	0.06183	6.6356	0.01	1.1727	1.0388 to 1.3237
BMD Femur Neck (g/cm2)	-6.6336	2.4351	7.421	0.0064	0.0013	0.0000 to 0.1555
Age (years)	0.06863	0.02589	7.027	0.008	1.071	1,0180 to 1,1268

PWV, pulse wave velocity; BMD, bone mineral density; DF, degree of freedom; SE, standard error; OR, odds ratio and CI, confidence interval.

Table 3 Age-Matched Predictive Value of Pulse Wave Velocity for All-Cause Mortality Using Residual-Based Matching

Backward Logistic Regression (PWV, age matched)						
Number of events	14	9,86%	8 (cardiovascular events) + 6 (Femur neck fatal outcome)			
Number censored	128	90,14%				
Total number of cases	142	100,00%				
Null model -2 Log Likelihood	91.42					
Full model -2 Log Likelihood	66.864					
Chi-squared	24.578					
DF	1					
Significance level	P < 0.0001					
Variable	Coefficient	SE	Wald	P	OR	95% CI of OR
Regresion_residual (PWV)	1.01872	0.24836	16.8254	< 0.0001	2.7697	1.7022 to 4.5064
Variable	b	SE	Wald	P		
Constant	-3.08051	0.48629	40.1292	< 0.0001		

PWV, pulse wave velocity; DF, degree of freedom; SE, standard error; OR, odds ratio; CI, confidence interval.

Table 4 Age-Matched Predictive Value of Femoral Neck Bone Mineral Density for All-Cause Mortality Using Residual-Based Matching

Backward Logistic Regression (PWV, age matched)						
Number of events	14	9,86%	8 (cardiovascular events) + 6 (Femur neck fatal outcome)			
Number censored	128	90,14%				
Total number of cases	142	100,00%				
Null model -2 Log Likelihood	91.42					
Full model -2 Log Likelihood	90.389					
Chi-squared	1.052					
DF	1					
Significance level	P = 0.3050					
Variable	Coefficient	SE	Wald	P	OR	95% CI of OR
Regresion_residual (BMD)	-2.09225	2.06044	1.0311	0.3099	0.1234	0.0022 to 7.0022
Variable	b	SE	Wald	P		
Constant	-2.24695	0.29022	59.9427	< 0.0001		

BMD, bone mineral density; DF, degree of freedom; SE, standard error; OR, odds ratio; CI, confidence interval.

PWV was a significant predictor (P = 0.01, OR = 1.1727, 95% CI 1.0388–1.3237).

Residual-based matching adjusted for age confirmed PWV as a strong predictor of adverse outcomes (OR = 2.7679, 95% CI 1.7022–4.5064), while FN BMD was not significant after age adjustment (OR = 0.1234, 95% CI 0.0022–7.0022) (Tables 3 and 4).

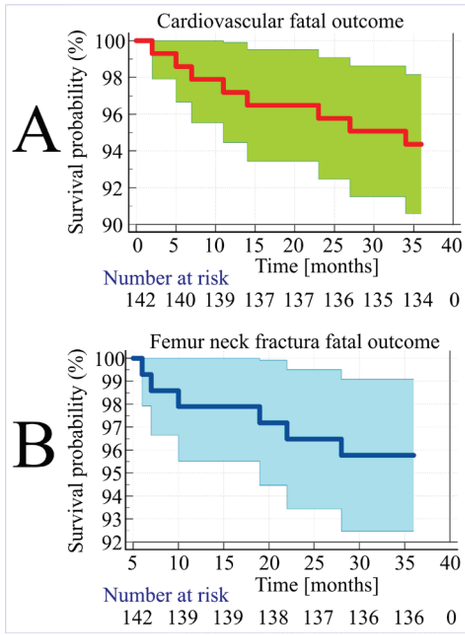


Figure 2 – Kaplan-Meier Survival Probability Curves for Fatal Outcomes Following Cardiovascular Events and Femoral Neck Fractures

Note: Each vertical decline represents a fatal event—8 in cardiovascular events and 6 in femoral neck fractures.

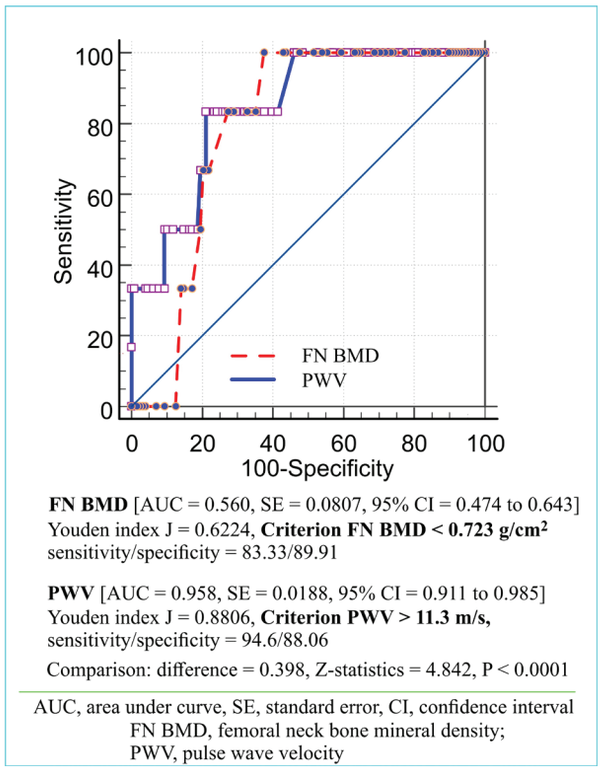


Figure 3 – Pairwise Comparison of Receiver Operating Characteristic Curves for Cardiovascular Death Predicted by Elevated Pulse Wave Velocity and Fatal Femoral Neck Fracture Predicted by Low Bone Mineral Density, Highlighting Significant Statistical Difference

Kaplan-Meier curves (Figure 2A, 2B) showed survival probabilities for cardiovascular and FN fracture deaths over 36 months. ROC analysis (Figure 3) demonstrated PWV had excellent discriminatory power for all-cause mortality [AUC = 0.985, J (Youden) index = 0.8806, cutoff 11.3 m/s, sensitivity 94.6%, specificity 88.06%], while FN BMD showed modest predictive ability (AUC = 0.560, J index = 0.6224, cutoff 0.723 g/cm², sensitivity 83.33%, specificity 89.91%). The optimal cutoff value was determined using the J (Youden) index, a summary measure of diagnostic performance calculated as sensitivity plus specificity minus one, which identifies the point maximizing the test's overall accuracy. The difference in predictive capability was significant (Z = 4.842, P < 0.0001), highlighting PWV as a superior prognostic marker compared to FN BMD.

Discussion

This study, which examined 142 patients, highlights the intricate interplay between skeletal and vascular health in predicting long-term outcomes, including fractures and mortality due to cardiovascular causes during a three-year observation period. Deceased individual in both the cardiovascular death group and FN fatal outcome group displayed a substantially higher age than survivors (p= 0.0036 and p =0.001, respectively). This indicates that advancing age correlates with an increased likelihood of mortality in both conditions. Advanced age plays a crucial role in cardiovascular-related death, partly due to the long-term influence of conventional predisposing elements. Studies confirm that while age's influence on risk decreases at very advanced ages, it remains a crucial factor in both short- and long-term mortality [16].

No meaningful difference was observed in BMI between survivors and non-survivors in either the cardiovascular or FN outcome groups (p =0.764 and p = 0.156, correspondingly), indicating that BMI may not be a predictor of mortality in these contexts. A notable difference was observed in smoking status between individuals who survived and those who did not in the cardiovascular group (p = 0.005), implying a conceivable nexus between tobacco consumption and lethal cardiovascular outcomes. Smoking significantly amplifies the likelihood of cardiovascular death, with a risk of coronary artery dysfunction that is two to four times higher and an excess mortality rate exceeding 70%, compounded by synergistic effects with traditional risk factors like hypertension, hypercholesterolemia, and diabetes. Nicotine and carbon monoxide exacerbate cardiovascular damage by impairing the myocardial oxygen supply/demand ratio, causing endothelial injury, and promoting atherosclerotic plaque development [17]. However, no substantial variation was identified in the FN group (p = 0.075).

Hypertension status alone may not distinguish survival outcomes in both groups. Non-survivors in both groups had lower BMD values compared to survivors, with a notable difference in the FN group (p = 0.0013). This suggests that lower BMD may be linked to poorer outcomes, particularly in FN fractures. PWV was markedly elevated in non-survivors compared to survivors only in cardiovascular estimated group (P = 0.0041), indicating that greater arterial stiffness correlates with an elevated risk of mortality [2, 11, 18].

In summary, age, smoking status, and clinical biomarkers (BMD and PWV) emerge as important factors associated with mortality in both cardiovascular and FN fracture outcomes.

The findings of this research underscore the critical role of vascular rigidity, quantified through PWV, in predicting

overall mortality, surpassing the predictive value of bone strength as measured by BMD at the FN. PWV emerged as a robust predictor, with each 1 m/s increase linked to 17.27% higher likelihood of mortality. This finding aligns with previous evidence linking increased arterial stiffness to cardiovascular events, systemic inflammation, and mortality [19]. In contrast, while BMD demonstrated a protective effect—each 1 g/cm² increase reduced the odds of mortality by 99.87%—its predictive capacity appears secondary to that of arterial stiffness. This means that low, osteopenic, or osteoporotic bone, with each 1 g/cm² decrease in BMD, demonstrated a higher risk of mortality, highlighting the significant impact of bone density on health outcomes.

The observed interplay between arterial stiffness and bone strength may be mechanistically linked through calcium and phosphate metabolism [20]. Heightened oxidative imbalance plays a pivotal role in arterial mineral deposition, facilitating bone-like transformation and hardening of vascular cells in conditions like diabetes, atherosclerosis, and chronic kidney disease [21]. Additionally, oxidative stress-driven molecular cascades within vascular smooth muscle cells exacerbate arterial rigidity and disease progression [22]. Together with dysregulated bone remodeling and increased circulating calcium, oxidative stress contributes to the deposition of hydroxyapatite crystals in arterial walls, linking vascular calcification to both arterial stiffness and bone strength [20, 21, 22]. Bone remodeling, a natural process involving the resorption of aged bone and the deposition of newly synthesized bone matrix, becomes disrupted with aging and chronic diseases. Increased bone resorption can lead to a shift of calcium from bone into the bloodstream. Elevated circulating calcium levels contribute to vascular calcification, a hallmark of arterial stiffness [23]. This pathological process involves the deposition of hydroxyapatite crystals, typically found in bone, into the arterial walls, leading to loss of elasticity and increased vascular rigidity, heightened arterial stiffness, and a rise in PWV [24]. Advancing age leads to arterial stiffness, intimal thickening, and calcification, increasing cardiovascular risk. In postmenopausal women, declining estrogen exacerbates these changes by reducing its protective effects on endothelial function and inflammation. This accelerates vascular aging, promoting plaque development and calcification, highlighting the need for targeted strategies in this population [24].

Concurrently, bone weakening occurs due to reduced bone mineralization and structural deterioration. Factors such as declining estrogen levels in postmenopausal women, chronic inflammation, and oxidative stress exacerbate bone loss, further amplifying the mineral imbalance [23, 25]. The findings underscore that while both BMD and PWV are significant predictors, PWV emerges as a stronger and more independent marker of all-cause and cardiovascular mortality. Age reflects cumulative biological changes and exposure to risk factors, with a consistent, moderate effect on mortality risk. Its OR of 1.071 and narrower confidence interval, compared to BMD, suggests stability in its association with mortality. BMD has a stronger association, as indicated by the extreme OR, but its wider confidence interval makes its predictive value less precise, possibly due to sample size or measurement variability. PWV, a robust indicator of arterial stiffness and cardiovascular risk, shows a slightly stronger effect on mortality than age, but its overall predictiveness is less stable due to a higher P-value and wider CI. The results reveal that lower FN BMD is closely linked to a higher likelihood of fractures and mortality, reflecting the profound impact of skeletal fragility on patient outcomes [2, 26]. However, the predictive power of BMD, though statistically significant, is less precise, likely due to broader confidence intervals and a potential overlap with other variables, such as age [27].

After adjusting for age using residual-based matching, our findings indicate that loss of arterial elasticity, as measured by PWV, plays a more prominent role in predicting fatal events compared to bone health. The strong association between increased PWV and higher event risk emphasizes the importance of arterial stiffness in cardiovascular mortality. In contrast, BMD did not demonstrate a significant influence on the likelihood of fatal events, indicating that factors other than bone density might have a greater impact on mortality outcomes in this population. These results highlight the need to focus on arterial health, particularly in aging individuals, while considering that bone health may not be as critical a factor in fatal events when age is controlled.

In contrast, PWV, a direct indicator of vascular stiffness, demonstrates robust predictive power for both cardiovascular death and all-cause mortality [2, 27]. The results of Vlachopoulos et al. (2010) align with our findings [27]. They reported an increased risk for total cardiovascular events (RR: 2.26), CV mortality [relative risk (RR): 2.02], and all-cause mortality (RR: 1.90) in individuals with high PWV, with a 14-15% rise in the likelihood of adverse outcomes for every 1 m/s elevation in PWV. An RR > 1 indicates an increased risk of the outcome in the exposed group compared to the reference group. Similarly, our study found that each unit increase in regression residual of PWV was correlated with significantly higher odds of CV events (OR: 2.77, $p < 0.0001$). The OR of 2.77 (95% CI 1.70–4.51) indicates that higher PWV is significantly associated with increased odds of the outcome, with a 95% probability that the true effect lies within this range. Both studies confirm that elevated PWV is a significant determinant of cardiovascular outcomes, emphasizing its prognostic value [27].

Our findings closely align with Khoshdel et al. (2007), who reported that a one-level increment in arterial PWV correlated with a mortality RR of 2.41 with 95% CI ranging from 1.81 to 3.20 and a cardiovascular event RR of 1.69 (95% CI: 1.35–2.11). They also observed significant differences in PWV between survivors and non-survivors across populations with both low and high risk highlighting its utility in cardiovascular risk stratification [28]. The consistency between our results and those of Khoshdel et al. underscores the potential of PWV as a valuable tool for assessing systemic vascular health and guiding patient management.

The association remains consistent, with PWV showing an incremental risk increase per unit rise, independent of age and other confounders. Given the strong correlation between age and arterial stiffness, the higher reliability and specificity of PWV suggest that it captures the critical vascular contributions to aging-related mortality better than age alone [29].

Similar to the findings of Meaume et al. (2001), which identified aortic PWV as a robust and independent indicator of cardiovascular mortality in individuals aged over 70 years, our study highlights that PWV, with its higher reliability and specificity, better captures the critical vascular contributions to aging-related mortality than age alone [29].

This suggests that PWV provides a more precise measure of the vascular aging process, offering superior predictive value for cardiovascular outcomes. While bone fragility elevates the likelihood of fractures and death, increased arterial stiffness, measured by PWV, silently reflects the long-term impact of

vascular aging on cardiovascular health [2, 30, 31].

Our findings indicate that, while bone loss and cardiovascular conditions often coexist in the aging population [30], the precise cause-and-effect connection between vascular abnormalities and skeletal health remains uncertain. One possible explanation is that increased mobilization of calcium from bones undergoing osteoporotic changes leads to enhanced transfer and deposition onto vascular walls, creating calcifications in large blood vessels (e.g., the aorta) and thereby increasing their stiffness. Additionally, the dysregulation of bone remodeling during osteoporosis, defined by a disruption in the balance between bone resorption and formation, may contribute to altered calcium metabolism and vascular calcification. Further research is needed to elucidate these complex interactions and their clinical implications. Although associations between increased PWV and lower BMD were observed, these relationships appeared to diminish after accounting for confounding factors.

The interplay between vascular and bone health may be partly explained by shared processes such as bone turnover and vascular calcification, which together form a key pathway linking reduced BMD with increased arterial stiffness. During accelerated bone resorption, calcium and phosphate are released into the circulation, promoting deposition of calcium in the vascular wall and contributing to medial arterial calcification. This process reduces arterial elasticity and elevates PWV, reflecting a shared pathophysiological pathway between skeletal demineralization and vascular aging. Moreover, osteoporosis-related bone loss enhances circulating calcium-phosphate product levels, further accelerating vascular calcification and stiffening [10, 12]. Increased oxidative stress, disrupted calcium and phosphate homeostasis, and hyperglycemia can contribute to vascular calcification while impairing bone mineralization [32]. Additionally, inflammatory signaling molecules including interleukin-1β cytokine and necrosis-inducing factor increase the expression of a crucial osteoclast-activating ligand, contributing to bone loss and vascular calcification [33]. Furthermore, decreased estrogen levels, particularly in postmenopausal women, may contribute to both osteoporosis and arterial stiffness [34]. These shared pathways underscore the complex bidirectional relationship between vascular and bone health. This suggests that impaired vascular function might not be the main pathway connecting osteoporosis and cardiovascular disease. However, akin to the research conducted by Ruicong et al. (2024), which examined the intricate relationship linking vascular stiffness and bone health, aging is not merely a passage of time but a process marked by the interconnected decline of both bone and vascular systems, requiring further research to clarify these complex interactions [30].

Both stiff arteries and weak bones pose significant risks as we age, but their impact varies. Stiff arteries can lead to life-threatening cardiovascular events like heart attacks and strokes, while weak bones increase the risk of fractures, leading to disability and loss of independence. While both are silent threats, the immediate danger of arterial stiffness often outweighs the long-term complications of bone fragility, making cardiovascular health a critical focus in aging. This study highlights a significant inverse relationship between PWV and BMD, suggesting that greater arterial stiffness, reflected by higher PWV, is associated with lower BMD, thereby linking vascular aging with skeletal fragility. The inverse relationship observed between PWV and BMD further reflects this bidirectional interaction, where increasing arterial stiffness parallels declining bone density, underscoring a shared pathophysiological pathway between vascular and skeletal deterioration.

Conclusions

PWV reflects the cumulative effect of vascular aging, providing actionable insights for risk stratification and management of age-related conditions. Elevated PWV is a powerful indicator of cardiovascular risk, with each unit increase in its regression residual linked to a 2.77-fold higher likelihood of such events, underscoring its prognostic value. While BMD retains importance in predicting fracture outcomes, its predictive value is secondary to PWV, which offers a more direct marker for mortality risk. PWV demonstrates high sensitivity and good specificity in predicting all-cause mortality, outperforming FN BMD as a prognostic marker for mortality risk in this population.

The inverse relationship between PWV and BMD suggests shared pathophysiological pathways, underscoring the interconnected nature of vascular and skeletal health. These findings advocate for the integration of PWV measurement in clinical practice to complement osteoporosis management strategies, forming a holistic approach to mitigating risks associated with aging. Future research into the molecular mechanisms linking vascular calcification and bone resorption could further enhance our understanding and enable targeted therapeutic interventions.

Limitations of the study

One major limitation of this research is the limited cohort of 142 participants also the short follow-up duration of 36 months, during which only 8 fatal cardiovascular events and 6 FN fracture-related deaths were observed. These constraints may limit the generalizability and statistical power of the findings. Future research should include larger cohorts with follow-up periods exceeding five years to capture a higher number of fatal events, both cardiovascular and fracture-related, for more robust and comprehensive conclusions.

Author Contributions: *M. A.,* the lead physician, oversaw participant selection, study implementation, and medical history analysis. She identified pathophysiological links between bone health and arterial stiffness, supported findings with references, ensured data analysis aligned with results, and critically reviewed the manuscript for scientific rigor. She also conducted a comprehensive literature review and played a key role in the study's conception, data analysis, and discussion. *P. A.* contributed to data organization, statistical validation, and ensuring methodological accuracy. He assisted in refining the discussion by integrating relevant findings and enhancing the manuscript's clarity and coherence. *L. T.,* contributed by analyzing diagnostic imaging, interpreting data, and ensuring accurate result interpretation. She also managed data organization in Excel, enhancing the clarity and reliability of the findings. *B. T.* contributed by sourcing relevant literature, refining the discussion, enhancing tables, and improving language, spelling, and grammar. *K. S.* applied statistical methods, interpreted results, and provided key insights, ensuring a strong data-driven foundation for the study. *D. Z.,* an informatics and cloud expert, managed data collection, storage, and processing. He supervised statistical methods and contributed to result interpretation in the discussion. All authors collaborated in writing, reviewing, and reaching a unanimous consensus on the final manuscript.

Disclosures: There is no conflict of interest for all authors.

Acknowledgments: None.

Funding: None.

Data availability statement: The data supporting the findings of this clinical study are included within the manuscript. Due to the sensitive nature of patient's information, additional data will not be made publicly available to maintain patient confidentiality. Specific data requests will be evaluated on a case-by-case basis, with consideration of ethical and privacy requirements.

Patient Informed Consent Statement: Written informed consent was obtained from all participants involved in the study, ensuring they understood the study's purpose, procedures, and their right to confidentiality.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. Guo J, Huang X, Dou L, Yan M, Shen T, Tang W, Li J. Aging and aging-related diseases: from molecular mechanisms to interventions and treatments. *Signal Transduct Target Ther.* 2022 Dec 16;7(1):391. https://doi.org/10.1038/s41392-022-01251-0
2. Avramovski P, Avramovska M, Sikole A. Bone Strength and Arterial Stiffness Impact on Cardiovascular Mortality in a General Population. *J Osteoporos.* 2016:7030272. https://onlinelibrary.wiley.com/doi/10.1155/2016/7030272.
3. Cheng CH, Chen LR, Chen KH. Osteoporosis Due to Hormone Imbalance: An Overview of the Effects of Estrogen Deficiency and Glucocorticoid Overuse on Bone Turnover. *Int J Mol Sci.* 2022;23(3):1376. https://www.mdpi.com/1422-0067/23/3/1376
4. Martini N, Streckwall L, McCarthy AD. Osteoporosis and vascular calcifications. *Endocr Connect*. 2023 Oct 5;12(11):e230305. https://ec.bioscientifica.com/view/journals/ec/12/11/EC-23-0305.xml
5. Cannata-Andía JB, Carrillo-López N, Messina OD, Hamdy NAT, Panizo S, Ferrari SL, On Behalf Of The International Osteoporosis Foundation I of Working Group On Bone And Cardiovascular Diseases. Pathophysiology of Vascular Calcification and Bone Loss: Linked Disorders of Ageing? *Nutrients*. 2021 Oct 27;13(11):3835. https://www.mdpi.com/2072-6643/13/11/3835
6. Tap L, Kirkham FA, Mattace-Raso F, Joly L, Rajkumar C, Benetos A. Unraveling the Links Underlying Arterial Stiffness, Bone Demineralization, and Muscle Loss. *Hypertension*. 2020 Sep;76(3):629-639. https://www.ahajournals.org/doi/10.1161/HYPERTENSIONAHA.120.15184
7. Vasan RS, Pan S, Xanthakis V, Beiser A, Larson MG, Seshadri S, Mitchell GF. Arterial Stiffness and Long-Term Risk of Health Outcomes: The Framingham Heart Study. *Hypertension.* 2022 May;79(5):1045-1056. https://www.ahajournals.org/doi/10.1161/HYPERTENSIONAHA.121.18776
8. Aibar-Almazán A, Voltes-Martínez A, Castellote-Caballero Y, Afanador-Restrepo DF, Carcelén-Fraile MDC, López-Ruiz E. Current Status of the Diagnosis and Management of Osteoporosis. *Int J Mol Sci.* 2022 Aug 21;23(16):9465. https://pubmed.ncbi.nlm.nih.gov/36012730/
9. Y, Han X, Liu J, Li Y, Li Z, Zhang W, Lv N, Dang A. Prognostic significance of the estimated pulse wave velocity in critically ill patients with coronary heart disease: analysis from the MIMIC IV database. *Eur Heart J Qual Care Clin Outcomes.* 2024 Sep 14:qcae076. https://academic.oup.com/ehjqcco/article/11/6/739/7758232?login=false
10. Avramovski P, Sikole A. The progression of bone mineral density loss in dialysis patients compared with the general population. *Korean J Intern Med*. 2012 Dec;27(4):436-42. https://kjim.org/journal/view.php?doi=10.3904/kjim.2012.27.4.436
11. Nikleski Z, Avramovska M, Avramovski P, Siklovska V, Zdraveski D, Trajcevska I, Stefanovska S, Spaseva-Karanfilova B, Aleksoska E, Sotiroski K. Arterial Stiffness as a Superior Predictor of Cardiovascular Risk: Comparative Analysis in Osteoporotic, Rheumatoid Arthritis, and Chronic Kidney Disease Populations. *Acta Scientific Orthopaedics.* 2024; 7: 5-7.
12. Avramovski P, Janakievska P, Sotiroski K, Sikole A. Accelerated progression of arterial stiffness in dialysis patients compared with the general population. *Korean J Intern Med*. 2013 Jul;28(4):464-74. https://www.kjim.org/journal/view.php?doi=10.3904/kjim.2013.28.4.464
13. Mulas O, Sestu A, Costa A, Chessa S, Vargiu C, Corda L, Pittau F, La Nasa G, Caocci G, Scuteri A. Arterial Stiffness as a New Predictor of Clinical Outcome in Patients with Polycythemia Vera. *J Clin Med.* 2024 Nov 13;13(22):6811. https://www.mdpi.com/2077-0383/13/22/6811
14. Pereira, T., Correia, C. & Cardoso, J. Novel Methods for Pulse Wave Velocity Measurement. *J. Med. Biol. Eng.* 2015; 35, 555–565. https://link.springer.com/article/10.1007/s40846-015-0086-8
15. Pilz N, Heinz V, Ax T, Fesseler L, Patzak A, Bothe TL. Pulse Wave Velocity: Methodology, Clinical Applications, and Interplay with Heart Rate Variability. *Rev Cardiovasc Med.* 2024; 25(7): 266. https://www.sciencedirect.com/science/article/abs/pii/S0025712511001192?via%3Dihub
16. Dhingra R, Vasan RS. Age as a risk factor. *Med Clin North Am.* 2012 Jan;96(1):87-91.
17. Lakier JB. Smoking and cardiovascular disease. *Am J Med.* 1992 Jul 15;93(1A):8S-12S. https://linkinghub.elsevier.com/retrieve/pii/000293439290620Q
18. Chen C, Bao W, Chen C, Wang L, Gong H. Association between estimated pulse wave velocity and all-cause mortality in patients with coronary artery disease: a cohort study from NHANES 2005–2008. *BMC Cardiovasc Disord.* 23, 412 (2023). https://doi.org/10.1186/s12872-023-03435-0
19. Mozos I, Malainer C, Horbańczuk J, Gug C, Stoian D, Luca CT, Atanasov AG. Inflammatory Markers for Arterial Stiffness in Cardiovascular Diseases. *Front Immunol.* 2017 Aug 31;8:1058. https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2017.01058/full
20. Chen Y, Zhao X, Wu H. Arterial Stiffness: A Focus on Vascular Calcification and Its Link to Bone Mineralization. *Arterioscler Thromb Vasc Biol.* 2020 May;40(5):1078-1093. https://www.ahajournals.org/doi/10.1161/ATVBAHA.120.313131
21. Nguyen NT, Nguyen TT, Da Ly D, Xia JB, Qi XF, Lee IK, Cha SK, Park KS. Oxidative stress by Ca2+ overload is critical for phosphate-induced vascular calcification. *Am J Physiol Heart Circ Physiol.* 2020 Dec 1;319(6):H1302-H1312. https://journals.physiology.org/doi/full/10.1152/ajpheart.00305.2020

22. De Paula, F., Rosen, C. Bone Remodeling and Energy Metabolism: New Perspectives. *Bone Res.* 2013; (1) 72–84. https://doi.org/10.4248/BR201301005
23. Cecelja M, Jiang B, Bevan L, Frost ML, Spector TD, Chowienczyk PJ. Arterial stiffening relates to arterial calcification but not to noncalcified atheroma in women. A twin study. *J Am Coll Cardiol.* 2011 Mar 29;57(13):1480-6. https://www.jacc.org/doi/10.1016/j.jacc.2010.09.079
24. Tesauro M, Mauriello A, Rovella V, Annicchiarico-Petruzzelli M, Cardillo C, Melino G, Di Daniele N. Arterial ageing: from endothelial dysfunction to vascular calcification. *J Intern Med.* 2017 May;281(5):471-482. https://onlinelibrary.wiley.com/doi/10.1111/joim.12605
25. Novella S, Heras M, Hermenegildo C, Dantas AP. Effects of estrogen on vascular inflammation: a matter of timing. *Arterioscler Thromb Vasc Biol.* 2012 Aug;32(8):2035-42. https://www.ahajournals.org/doi/10.1161/ATVBAHA.112.250308
26. Li HL, Shen Y, Tan LH. Fu SB, Dai RC, Yuan LQ, Sheng ZF, Xie ZJ, Wu XP, Liao EY, Tang XL, Wu XY. Relationship between bone mineral density and fragility fracture risk: a case-control study in Changsha, China. *BMC Musculoskelet Disord.* 22, 728 (2021). https://doi.org/10.1186/s12891-021-04616-8
27. Vlachopoulos C, Aznaouridis K, Stefanadis C. Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-analysis. *J Am Coll Cardiol.* 2010 Mar 30;55(13):1318-27. https://www.jacc.org/doi/10.1016/j.jacc.2009.10.061
28. Khoshdel AR, Carney SL, Nair BR, Gillies A. Better management of cardiovascular diseases by pulse wave velocity: combining clinical practice with clinical research using evidence-based medicine. *Clin Med Res.* 2007 Mar;5(1):45-52. https://www.clinmedres.org/content/5/1/45
29. Meaume S, Benetos A, Henry OF, Rudnichi A, Safar ME. Aortic pulse wave velocity predicts cardiovascular mortality in subjects >70 years of age. *Arterioscler Thromb Vasc Biol.* 2001 Dec;21(12):2046-50. 11742883. https://www.ahajournals.org/doi/10.1161/hq1201.100226
30. Xue R, Zhang J, Zhen Z, Liang W, Li Y, Zhang L, Dong Y, Dong B, Liu C. Estimated pulse wave velocity predicts mortality in patients with heart failure with preserved ejection fraction. *Hellenic J Cardiol.* 2024 May 23:S1109-9666(24)00117-9. https://www.ahajournals.org/doi/10.1161/hq1201.100226
31. Usiskin IM, Mitchell GF, Bouxsein ML, Liu CT, Kiel DP, Samelson EJ. Vascular function and skeletal fragility: a study of tonometry, brachial hemodynamics, and bone microarchitecture. *J Bone Miner Res.* 2024 Aug 5;39(7):906-917. https://doi.org/10.1093/jbmr/zjae071
32. Chen Y, Zhao X, Wu H. Arterial Stiffness: A Focus on Vascular Calcification and Its Link to Bone Mineralization. *Arterioscler Thromb Vasc Biol.* 2020 May;40(5):1078-1093. https://www.ahajournals.org/doi/10.1161/ATVBAHA.120.313131
33. Zhang M, Bai L, Kang J, Ge J, Peng W. Links between arterial stiffness and bone mineral density in middle-aged and elderly Chinese individuals: a cross-sectional study. *BMJ Open.* 2019 Aug 10;9(8):e029946. https://bmjopen.bmj.com/content/9/8/e029946
34. Wang YQ, Yang PT, Yuan H, Cao X, Zhu XL, Xu G, Mo ZH, Chen ZH. Low bone mineral density is associated with increased arterial stiffness in participants of a health records based study. *J Thorac Dis.* 2015 May;7(5):790-798. https://doi.org/10.3978/j.issn.2072-1439.2015.04.47

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17358

Health Care Quality and Patient Satisfaction in Hemodialysis Clinic: Investigation of Patient-Reported Outcomes

Meltem Saygili[1], Aysun Kandemir Ture[1], Nesrin Akca[1]

[1]*Department of Health Management, Faculty of Health Sciences, Kirikkale University, Turkey*

Received: 2025-09-08.
Accepted: 2025-10-06.

J Clin Med Kaz 2025; 22(6): 19–27

Corresponding author:
Meltem Saygili
E-mail: meltemsaygili@kku.edu.tr.
ORCID: 0000-0001-6309-2473

ABSTRACT

Background: Hemodialysis treatment requires a long-term and continuous relationship between the patient and the healthcare institution. Therefore, patient evaluations regarding the services provided are highly valuable. This study was conducted to assess patients' perceptions of the quality of services offered at the hospital and to examine the relationship between these perceptions and patient satisfaction.

Methods: This cross-sectional study included adult patients receiving hemodialysis treatment at a public hospital in Türkiye (n = 105). Data was collected using a face-to-face survey method. The survey included an information form about patients' sociodemographic and clinical characteristics, the SERVQUAL Scale, and the Patient Satisfaction Questionnaire (PSQ-18). The data were evaluated using descriptive statistics, nonparametric tests (Mann-Whitney U test and Kruskal-Wallis test), and Spearman correlation analysis. Additionally, the gap scores between expected and perceived service quality were calculated for the SERVQUAL Scale.

Results: A negative gap was identified between participants' quality expectation scores and their perceived service quality scores (SERVQUAL = -0.26). The most significant differences were found in the Reliability (-0.43), Tangibles (-0.38), and Assurance (-0.35) sub-dimensions. Participants' patient satisfaction scores were at a medium level (median: 54.00, min: 41, max: 71). Correlation analysis results indicated a statistically significant moderate positive relationship between patients' service quality scores and patient satisfaction levels ($r = 0.501$, $p < 0.01$) in the hemodialysis unit.

Conclusion: The results reveal that the quality expectations of hemodialysis patients are not adequately met and that patient satisfaction is not high. Furthermore, it has been found that the perception of service quality is positively and moderately related to patient satisfaction. The findings are expected to guide healthcare institutions and managers in providing patient-centered care, delivering quality services, and improving patient satisfaction for this specific patient group.

Keywords: Perceived service quality, Expected service quality, Patient satisfaction, Hemodialysis Patients, Chronic Kidney Diseases.

Introduction

Chronic kidney disease (CKD) represents a significant global health challenge, affecting approximately 10% of the world's population and contributing to substantial morbidity and mortality [1]. The prevalence of CKD is expected to rise, driven by increasing life expectancy and the growing incidence of diabetes and hypertension, the two leading causes of CKD [2-5]. The chances of survival and quality of life of patients with end-stage kidney disease (ESKD) are generally maintained by renal replacement therapy, which includes hemodialysis,

peritoneal dialysis, and kidney transplantation [6]. For eligible patients, transplantation offers a better quality of life, more prolonged survival, and lower costs compared to dialysis [7]. However, most patients opt for dialysis treatment due to infrastructure deficiencies, a few donors, and contraindications concerning transplantation. Even in countries that are more active in transplant operations, approximately 30-50% of the patients who need renal replacement therapy receive dialysis treatment [8]. More than 1 in 7 U.S. adults, approximately 35.5 million people, are estimated to have CKD, and approximately 808.000 people in the U.S. are living with End Stage Kidney Disease (ESKD), with 69% on dialysis and 31% having received a kidney transplant [9,10]. Therefore, hemodialysis is the most common treatment option referred to worldwide for the treatment of ESKD [12].

Hemodialysis is a repetitive and complex procedure for patients, often performed three times a week at short intervals in hospitals or dialysis centers. Therefore, patients must spend most of their lives in hospitals/hemodialysis centers [11]. Quality of service is a crucial factor that positively impacts patient satisfaction among kidney patients [13]. Studies have revealed that the expectations of patients receiving hemodialysis exceed their perceptions of the quality of the service provided [14,15] and that the requirements of hemodialysis patients are not fully met in healthcare institutions [16,17]. Hence, evaluating the expectations and perceptions of service quality among these patients, who must spend a considerable amount of time in the hospital, presents a significant opportunity to identify weaknesses in healthcare services and implement improvements in hemodialysis units.

Furthermore, Serrano-del Rosal and Loriente-Arín argued that the satisfaction of healthcare service beneficiaries significantly impacts their tendency to follow medical and therapeutic recommendations, thereby improving their health [18]. Additionally, it was emphasized that service redesign is necessary to meet patients' needs and that improving service delivery is essential to reduce the emotional and financial costs associated with hemodialysis [19]. Therefore, continuous efforts to improve the quality of healthcare services, particularly in hemodialysis, are essential. This includes addressing service quality gaps, enhancing patient-provider communication, and ensuring that care is patient-centered and empathetic.

This study was conducted with patients undergoing hemodialysis treatment using established models such as SERVQUAL and the Patient Satisfaction Questionnaire-18 (PSQ-18). Specifically, the study aimed to: (i) evaluate the service quality perceived and expected by patients, (ii) assess whether there is a gap between perceived and expected service quality, (iii) to examine the level of patient satisfaction, (iv) to evaluate whether service quality perception and patient satisfaction differ according to the sociodemographic characteristics of patients, and (v) to reveal whether there is a relationship between patients' perceptions of service quality and their satisfaction levels. By identifying areas for improvement, healthcare providers can better meet the needs of CKD patients, ultimately enhancing their quality of life and treatment outcomes. The results of this study are expected to provide evidence and guidance to healthcare providers, healthcare managers, and policymakers.

The research hypotheses determined in line with the above specific objectives are presented below.

H1: There is a gap between the expected and perceived service quality among hemodialysis patients.

H2: The perception of service quality among hemodialysis patients differs according to their sociodemographic characteristics.

H3: Patient satisfaction among hemodialysis patients differs according to their sociodemographic characteristics.

H4: The perception of service quality in hemodialysis patients is positively related to patient satisfaction.

Methods

Study design

This study was designed as a descriptive and cross-sectional study.

Study population

The study was conducted between January and February 2020 at a state hospital serving Kırıkkale, Türkiye, with the participation of patients with chronic kidney disease undergoing hemodialysis treatment. According to hospital records, during the study, a total of 122 patients were receiving hemodialysis treatment at the hospital's hemodialysis clinic on a rotating basis. The inclusion criteria for the study were: being conscious, being an adult, speaking Turkish, and agreeing to participate voluntarily. No sample was selected; all patients receiving hemodialysis treatment during the relevant period were invited to participate in the study. During the study, nine patients declined to participate, four patients could not speak Turkish, one patient died, and three patients were transferred to other treatment units. Data was collected using a face-to-face questionnaire method. After providing the necessary information and explanations, patients were asked to complete the questionnaire forms. For illiterate patients, the researcher read the questions aloud, and responses were obtained. The study was completed with 105 patients who agreed to participate voluntarily (participation rate: 86.06%).

Christian Grönroos first proposed the concept of perceived service quality, which expresses the three main determinants of service quality: technical quality, functional quality, and corporate image [20]. In 1985, Parasuraman et al. [21] developed the gap model, which is based on the Perceived Quality Model. In 1988, they simplified this model to establish a five-dimensional scale of service quality. The model is used to describe consumer-perceived gaps in service quality. The SERVQUAL scale comprises 22 items across five dimensions: physical appearance, reliability, responsiveness, assurance, and empathy. The dimensions of the SERVQUAL model are named as follows and address the specified elements [22]:

Tangibility: Appearance of physical facilities, equipment, employees, and communication materials.

Reliability: Ability to perform the promised service accurately with no errors

Responsiveness: The willingness to respond to and meet customer needs promptly and efficiently, as well as the attitude of the employees.

Assurance: Knowledge and the ability to provide information about the service provided, courtesy of employees, and their ability to inspire trust and confidence

Empathy: Individualized customer service.

The scale measures the quality of the service delivered by assessing the gaps between the expectations before delivery and the experience gained from using the service [23]. When gap scores on the SERVQUAL scale are greater than 0, it is concluded that the healthcare institution in question provides healthcare

services of a quality that exceeds patient expectations; when the score is less than 0, it is concluded that patient expectations have not been met. [21].

The Patient Satisfaction Questionnaire Short-Form (PSQ-18) is a short-form version of the 50-item PSQ-III. The PSQ-18 measures patients' satisfaction with medical services across seven dimensions, comprising 18 closed-ended questions: general satisfaction, technical quality, interpersonal manner, communication, financial aspect, time spent with the doctor, and accessibility and compliance [24]. All items on the scale were rated on a five-point Likert-type scale, with responses ranging from 1 (strongly agree) to 5 (strongly disagree). The service satisfaction score, as measured by the PSQ-18, was used in this study. According to the PSQ-18 scoring system, the total score for all sub-scales ranges between 18 and 90; 18 points represent the lowest possible rating, and 90 points represent the highest. Higher scores indicate higher patient satisfaction with the service provider [25]. Cronbach's alpha coefficient was calculated to assess the scale's internal consistency, indicating an acceptable level of reliability.

The question form, which includes the socio-demographic characteristics and clinical information of the patients, involves questions about the following: age, gender, marital status, employment status, education, average monthly income, the presence of other chronic diseases, dialysis time at that hospital, whether sufficient information about the disease was provided, whether the patient would prefer this hospital again in case of any health problem and whether the patient would recommend this hospital to other dialysis patients.

Statistical analysis

Descriptive statistics (frequency, percentage, etc.) were used to summarize participants' disease-related data and socio-demographic characteristics. The reliability of the scales was evaluated with Cronbach's alpha coefficient. Data analysis was performed using the Statistical Package for the Social Sciences (SPSS) version 24.0 for Windows and Microsoft Excel 2010. Non-parametric tests (Mann-Whitney U test and Kruskal-Wallis test) were used because the data did not exhibit a normal distribution. Spearman correlation analysis was also applied to examine the relationships between SERVQUAL and PSQ-18 scales. The following formula was used to calculate the gap score:

"Gap score = Perception score (P) – Expectation score (E)"

Negative scores obtained with the gap score indicate that customers perceive the quality of service they receive as low, while positive scores indicate a higher perception of service quality.

Results

It was found that 54.3% of the hemodialysis patients were male, 60% were over 60 years old, and most of them were married and unemployed (Table 1). 50.5% of the patients stated that they had been receiving dialysis treatment at this hospital for over two years. 61.9% of patients had one or more chronic diseases in addition to kidney disease. Most patients had information about kidney disease, stated that they would prefer the same hospital when they experience any health problems, and would recommend the hospital to other hemodialysis patients.

Table 1 Descriptive statistics of the participants

Descriptive characteristics	Mean ± SD (Range)	n	%
Gender			
Female		48	45.7
Male		57	54.3
Age (years)			
<60	61.0 ± 13.5	42	40.0
≥60	(26-83)	63	60.0
Marital status			
Married		91	86.7
Single		14	13.3
Education			
Illiterate		24	22.9
Only literate		10	9.5
Primary school		45	42.9
Secondary school and above		26	24.8
Working status			
Working		11	10.5
Not working		94	89.5
Income (n=103)			
Low		18	17.5
Medium		48	46.6
High		37	35.9
Dialysis time at that hospital (month)			
≤24	43.2 ± 45.9	52	49.5
>25	(1-240)	53	50.5
Presence of other chronic diseases			
Yes		65	61.9
No		40	38.1
Have you been informed about kidney disease?			
Yes		82	78.1
No		23	21.9
Would you prefer this hospital again in case of any health problems?			
Yes		97	92.4
No		8	7.6
Would you recommend the hospital to other dialysis patients?			
Yes		98	93.3
No		7	6.7

SD: Standard Deviation

Table 2 presents the reasons why patients have preferred this hospital for treatment. The patients stated that they preferred this hospital based on their confidence in the state hospital, the

Table 2 Descriptive statistics of the participants

Reasons	n*
Because I trust the public hospital	57
Accessible	38
Because the health personnel are concerned/smiling	21
Because I think I get better quality service	19
Because people recommended it, I know	8
Because my doctor works in this hospital	6
Because it has more modern tools and equipment	6
Because I think it is more hygienic	5
Insufficient financial situation	5
Obligation-referral	3

** Number in preferences.*

Table 3 SERVQUAL scores of participants

Dimensions	Items	Expectation Scores Mean ± SD	**Perception Scores Mean ± SD**	**SERVQUAL Scores**	SERVQUAL Scores by Dimensions	Overall Score
Tangibility	1	4.68 ± 0.67	4.12 ± 0.80	−0.55	−0.38	−0.26
	2	4.37 ± 0.90	4.09 ± 0.79	−0.28		
	3	4.46 ± 0.80	4.20 ± 0.74	−0.26		
	4	4.55 ± 0.71	4.11 ± 0.74	−0.44		
Reliability	5	4.75± 0.51	4.41 ± 0.79	−0.34	−0.43	
	6	4.75 ± 0.50	4.30 ± 0.77	−0.45		
	7	4.82 ± 0.39	4.47 ± 0.69	−0.35		
	8	4.77 ± 0.42	4.27 ± 0.71	−0.50		
	9	4.66 ± 0.68	4.13 ± 0.79	−0.52		
Responsiveness	10	4.44 ± 0.83	4.31 ± 0.74	−0.13	−0.05	
	11	4.22 ± 1.02	4.34 ± 0.71	0.12		
	12	4.47 ± 0.83	4.33 ± 0.82	−0.13		
	13	3.98 ± 1.09	3.92 ± 1.13	−0.06		
Assurance	14	4.61 ± 0.67	4.34 ± 0.77	−0.27	−0.35	
	15	4.70 ± 0.46	4.30 ± 0.73	−0.41		
	16	4.69 ± 0.47	4.47 ± 0.67	−0.22		
	17	4.54 ± 0.60	4.04 ± 0.80	−0.50		
Empathy	18	4.02 ± 1.14	3.79 ± 1.19	−0.23	−0.08	
	19	3.91 ± 1.25	3.75 ± 1.22	−0.16		
	20	4.13 ± 1.04	4.13 ± 0.99	0.00		
	21	4.02 ± 0.98	3.89 ± 1.06	−0.13		
	22	3.93 ± 1.22	4.08 ± 0.95	0.14		
Cronbach's Alpha		0.858	0.914	0.889		

SD: Standard Deviation

Table 4 Participants' satisfaction scores (PSQ-18)

Dimensions (Number of items)	Mean	SD	**Scores Mean**	**SD**	Median	Min	Max
General satisfaction (2)	3.74	1.03	5.97	1.19	6.00	3.00	10.00
Technical quality (4)	4.10	0.73	12.30	1.69	12.0	9,00	20.00
Interpersonal manner (2)	4.12	0.82	6.09	1.08	6.00	2.00	10.00
Communication (2)	3.10	0.85	6.42	1.28	6.00	4.00	10.00
Financial aspect (2)	3.83	0.93	5.69	1.51	6.00	2.00	9.00
Time spent with the doctor (2)	3.48	0.91	5.36	1.32	6.00	2.00	9.00
Accessibility and convenience (4)	4.01	0.64	12.57	1.99	13.00	5.00	18.00
PSQ overall score	3.93	0.53	54.39	5.20	54.00	41.00	71.00
Cronbach's Alpha (PSQ Total): 0.765							

SD: Standard Deviation

hospital's accessible location, the caring/smiling faces of the healthcare personnel, and the delivery of better-quality health services.

Table 3 presents the scores from the SERVQUAL scale, indicating patients' expectations and perceptions of the quality of health services they received. Cronbach's alpha value of the sub-scale with the expectation expressions in the SERVQUAL scale was found to be 0.858, whereas the reliability of the sub-scale with the perception expressions was found to be 0.914. General SERVQUAL items of the scale have an alpha value of 0.889. These results confirmed that the scale is reliable.

Service quality perceptions of patients regarding healthcare services were generally lower than their expectations in all sub-dimensions. The most prominent use of the SERVQUAL scale is to determine the quality gaps in service delivery (Table 3). Study findings indicated that the highest quality gap based on dimensions is in the reliability sub-dimension, whereas the lowest quality gap is in the responsiveness dimension. In other words, the patients' expectations regarding the reliability sub-dimension were higher; however, their expectations could not be met.

On the other hand, the patients' expectations regarding the responsiveness sub-dimension, which represents the desire to respond to and meet patients' needs quickly and efficiently, were low. However, the quality perception of the services delivered in the healthcare institution is low. Patients' evaluations of the hospital's overall service quality were generally negative (overall SERVQUAL score −0.26). In line with these results, it can be concluded that the perceived service quality of patients receiving hemodialysis service from the relevant institution was below their expectations. According to these results, hypothesis H1 is accepted.

Table 5 Participants' satisfaction scores (PSQ-18)

Socio-demographic Characteristics	Gender		Age		Level of income		
	Female	Male	<60	≥60	Low	Medium	High
Overall SERVQUAL	58.2	48.6	49.0	55.7	61.9	52.2	47.0
	u=1118.0 p=0.108		u=1156.0 p=0.275		χ2=3.011 p= 0.222		
Expectation	53.1	53.0	50.8	54.5	48.4	53.9	51.3
	u=1365.0 p=0.985		u=1230.5 p=0.545		χ2=0.478 p= 0.787		
Perception	60.3	46.9	44.5	58.7	56.9	55.0	45.7
	u=1020.0 p=0.025		**u=965.0 p=0.019**		χ2=2.589 p= 0.274		
Overall PSQ-18	64.3	43.5	45.4	58.1	70.5	49.5	46.2
	u=825.0 p<0.001		**u=1004.0 p=0.037**		**χ2=8.617 p= 0.013**		
Overall satisfaction	62.4	45.1	47.9	56.4	66.3	51.5	45.8
	u=919.0 p=0.003		u=1109.0 p=0.155		χ2=5.904 p= 0.052		
Technical quality	62.7	44.9	48.0	56.3	56.7	54.5	46.5
	u=904.5 p=0.002		u=1113.0 p=0.160		χ2=2.152 p= 0.341		
Interpersonal manners	60.3	46.8	45.5	58.0	65.9	50.4	47.4
	u=1016.0 p=0.020		**u=1007.5 p=0.034**		χ2=5.168 p= 0.075		
Communication	60.2	46.9	45.2	58.2	66.1	55.0	41.2
	u=1020.5 p=0.022		**u=993.5 p=0.027**		**χ2=9.781 p= 0.008**		
Cost	58.6	48.3	58.9	49.1	57.3	42.0	62.4
	u=1099.0 p=0.078		u=1077.5 p=0.102		**χ2=10.820 p= 0.004**		
Time spent with the doctor	58.2	48.6	44.7	58.5	65.3	49.4	48.9
	u=1119.0 p=0.104		u=974.5 p=0.021		χ2=4.440 p= 0.109		
Accessibility and compliance	62.2	45.3	47.8	56.5	69.6	50.8	45.1
	u=926.5 p=0.004		u=1104.5 p=0.150		**χ2=8.422 p= 0.015**		

Table 4 exhibits the satisfaction scores reported by patients receiving hemodialysis services. Cronbach's alpha value for the PSQ-18 is 0.765. The mean total satisfaction score was 54.39 ± 5.20. The results indicated that patient satisfaction scores were at a medium level (median: 54.00, min: 41–max: 71); the lowest satisfaction was measured in the communication sub-dimensions, and the highest was measured in the interpersonal manner and technical quality sub-dimensions.

Table 5 presents the findings comparing participants' scores from the SERVQUAL and PSQ-18 scales according to their socio-demographic characteristics. The results show that there is no statistically significant difference between the scores obtained from the SERVQUAL (Total) scale and the socio-demographic characteristics examined (age, gender, income level, marital status, education level, and employment status) ($p > 0.05$). While participants' SERVQUAL (Expected) scores did not differ according to socio-demographic characteristics, SERVQUAL (Perceived) scores showed significant differences according to gender and age. Accordingly, women and patients aged 60 and over were found to have higher quality perception scores compared to men and younger patients ($p < 0.05$). In other words, they perceive the quality of the services provided as good. (Table 5). According to these results, hypothesis H2 is accepted.

Overall patient satisfaction, technical quality, interpersonal manner, communication, accessibility, and compliance sub-dimensions of the PSQ-18, as well as the overall PSQ-18 score, were found to differ significantly according to gender ($p < 0.05$). Specifically, women's satisfaction scores in these sub-dimensions were higher than those of men. Interpersonal manners, communication, time spent with the doctor, and overall PSQ-18 score differed significantly according to age ($p < 0.05$). Accordingly, participants aged 60 and over were found to have higher satisfaction scores in these dimensions compared to those under 60 years of age. On the other hand, statistically significant differences were observed in communication, financial dimension, accessibility, and compliance, and overall PSQ-18 scores, when measured according to income level ($p < 0.05$). The results show that low-income patients scored higher satisfaction scores than patients with higher incomes (Table 5). According to these results, hypothesis H3 is accepted.

Table 6 (see it on the next page) presents the results of the Spearman correlation analysis between SERVQUAL and PSQ-18 scale scores, as well as their subdimensions. Upon examining Table 6, it is observed that a positive and significant relationship exists between perceived service quality and patient satisfaction among hemodialysis patients ($r = 0.501$, $p < 0.01$). The results indicate that perceived service quality is positively correlated with patient satisfaction, particularly in terms of reliability ($r = 0.448$, $p < 0.01$) and responsiveness ($r = 0.310$, $p < 0.01$), exhibiting a moderate and statistically significant relationship. In other words, if the services provided to these patients in the institution wholly and accurately meet their needs, and healthcare professionals respond to these needs promptly, patient satisfaction levels can be increased. According to these results, hypothesis H4 is accepted.

On the other hand, it was found that the sub-dimensions of the Patient Satisfaction Scale, "time spent with the doctor" ($r = 0.326$, $p< 0.01$) and "accessibility" ($r = 0.432$, $p< 0.01$), were positively related to the patients' SERVQUAL scale overall scores at a moderate level. In summary, if the time spent with the doctor and the accessibility of services and healthcare personnel are improved, patients' perceptions of service quality may increase.

Discussion

The results obtained generally revealed that patients' perceptions of service quality regarding healthcare services were lower than their expectations. The highest quality gap was observed in the reliability sub-dimension, and the lowest quality

Table 6 Relationships between participants' quality expectations/perceptions, and patient satisfaction

Parameter		Tangibility	Reliability	Assurance	Responsiveness	Empathy	Expectation	Perception	SERVQUAL
Overall satisfaction	r	0.217*	0.445**	0.176	0.261**	0.129	0.069	0.436**	0.329**
	p	0.026	0.000	0.073	0.007	0.191	0.482	0.000	0.001
Technical quality	r	0.113	0.427**	0.220*	0.186	0.045	0.187	0.434**	0.208*
	p	0.250	0.000	0.024	0.058	0.649	0.056	0.000	0.033
Interpersonal	r	0.187	0.251**	0.192	0.228*	0.112	0.105	0.359**	0.252**
	p	0.056	0.010	0.050	0.019	0.256	0.286	0.000	0.009
Communication	r	0.149	0.348**	0.303**	0.176	0.010	0.211*	0.424**	0.202*
	p	0.130	0.000	0.002	0.073	0.923	0.031	0.000	0.039
Cost	r	−0.239*	−0.095	−0.127	−0.044	−0.090	0.029	−0.177	−0.138
	p	0.014	0.333	0.196	0.659	0.360	0.773	0.071	0.161
Time spent with the doctor	r	0.148	0.276**	0.317**	0.257**	0.221*	0.091	0.458**	0.326**
	p	0.131	0.004	0.001	0.008	0.023	0.353	0.000	0.001
Accessibility	r	0.284**	0.403**	0.220*	0.295**	0.271**	−0.040	0.386**	0.432**
	p	0.003	0.000	0.024	0.002	0.005	0.682	0.000	0.000
PSQ-18	r	0.198*	0.448**	0.270**	0.310**	0.154	0.137	0.501**	0.361**
	p	0.043	0.000	0.005	0.001	0.116	0.162	0.000	0.000

**p < 0.05 level (two-tailed).*
***p < 0.01 level (two-tailed).*

gap was in the responsiveness sub-dimension. Other studies in the literature support these findings. For instance, Bahadori et al. determined, in their study conducted with individuals with CKD, that the service quality perceptions of the patients regarding the health care services were lower than their expectations. It was further determined in the research mentioned above that the highest mean service quality perception score was obtained in the assurance sub-dimension (4.30 ± 0.36), and the lowest mean service quality perception score was obtained in the empathy sub-dimension (3.84 ± 0.34) [14]. Bowen Bucheli and Fosado Téllez determined that quality perception is high only in the sub-dimension of empathy [16]. In another study conducted in Türkiye, it was demonstrated that the quality of service perceived by CKD patients was most influenced by the "tangibles" subdimension and least influenced by the "empathy" subdimension [15].

This study revealed that women and patients aged 60 and over have higher quality perception scores than men regarding the care service delivered ($p < 0.05$). It can be concluded that young male patients have higher expectations and require services in more technologically advanced and professional environments. This result is similar to the literature [26].

Our study has revealed that patients are moderately satisfied with the services. In particular, it has been determined that the satisfaction levels of patients aged 60 and over, women, and those with low income are statistically significantly higher than those of other groups ($p < 0.05$). This finding is consistent with the existing literature [27, 28]. On the contrary, a study conducted with CKD patients in Türkiye found that male patients had higher satisfaction levels than female patients [29].

This research was conducted in the hemodialysis clinic of a public hospital providing services in Türkiye. Most of the patients who participated in the study (60%) were aged 60 years or older. The majority had been receiving treatment at the same hospital for a long time (24 months or more). Most participants stated that they would choose the same hospital for treatment again (92.4%) and would recommend it to others (93.3%). This can be attributed to the trust in public hospitals in Turkey. The highest satisfaction scores were obtained in the "interpersonal communication and technical quality" sub-dimensions.

In contrast, the lowest satisfaction scores were obtained in the "communication" sub-dimension. Additionally, it was found that low-income patients scored higher than patients with higher incomes in the sub-dimensions of communication, financial dimension, accessibility, and adaptation, as well as in the total PSQ-18 scale ($p < 0.05$). The lower expectations of patients with lower income levels can explain this finding.

Finally, the results obtained in this study revealed a significant positive correlation between "perceived service quality" and patient satisfaction ($r = 0.501$, $p < 0.01$). The most crucial quality perception sub-dimensions affecting the satisfaction of hemodialysis patients were found to be "reliability" and "responsiveness". Similarly, another study conducted in India revealed that the sub-dimensions of "reliability" and "responsiveness" of service quality perception affect patient satisfaction [30]. The patients' total SERVQUAL scale scores were found to be moderately positively correlated with "time spent with the doctor ($r = 0.326$, $p < 0.01$)" and "accessibility ($r = 0.432$, $p < 0.01$)" sub-dimensions of patient satisfaction. These results are supported by other studies conducted with CKD patients in national and international literature [17,31-37]. In these studies, the factors that most influenced patient satisfaction in CKD patients were ranked as concrete factors such as reliability, empathy, staff behavior, courtesy, persuasiveness, accessibility of services, accurate information, and communication [17,31-37]. Another study conducted in India revealed that accreditation statistically significantly improved the average satisfaction scores of hemodialysis patients compared to those before accreditation [38].

Conclusion

The results obtained from the study indicate that the perceptions of hemodialysis patients regarding the quality of healthcare services are lower than their expectations; in other words, there is a negative gap between the expected quality of services and the perceived quality of services. The most important contribution of this study to the literature is the identification of these gap areas. Indeed, the highest quality

difference between expected and perceived service quality was found in the “reliability (accurate and complete delivery of services)”, “tangibility (condition of physical facilities, equipment, staff, and communication materials)” assurance (Employees' ability to provide information about the services offered, their courtesy, and their ability to inspire confidence" dimensions. This suggests that hemodialysis patients who require hospital visits 2-3 times a week expect more modern devices and a more comfortable treatment environment. Furthermore, patient perceptions of reliability can be improved by enhancing the reliability of services, providing clear and effective information about the clinical process to healthcare personnel, implementing accurate and timely treatment interventions, and ensuring patient participation in clinical processes and decision-making. Furthermore, although the patients participating in the study generally achieved high satisfaction scores, they were less satisfied with “interpersonal approach” and “technical quality.” These results are consistent with patients' quality perception assessments.

Providing participant and patient-centered services is a common goal of all healthcare institutions. Feedback obtained directly from patients regarding the services provided in healthcare institutions offers valuable information for managers to inform future planning and identify improvements that need to be implemented. By addressing these areas, the hospital can work towards closing the gap between patient expectations and their perceptions of service quality, ultimately leading to higher overall patient satisfaction.

Limitations

This study is a single-center investigation conducted with patients receiving hemodialysis treatment at a state hospital in Türkiye; therefore, the results cannot be generalized to patients in other countries or regions. Additionally, patients' assessments of service quality and patient satisfaction are based on their statements, which may be considered a potential source of bias. Furthermore, the mandatory nature of hemodialysis services, which must be repeated at specific intervals, the necessity for patients to continue receiving this service until they undergo kidney transplantation or pass away, and the limited-service capacity of the institution have contributed to the limited number of patients included in the study. Finally, the survey method used to obtain research data is susceptible to social desirability bias. Including qualitative research methods in future studies may increase the depth of the analysis. We also recommend that researchers contribute to the literature through quality measurements and satisfaction assessments in centers of different sizes or through cross-cultural studies.

Author Contributions: Conceptualization, M. S. and A. K. T.; methodology, M. S., A. K. T. and N. A.; formal analysis, M. S. and A. K. T.; investigation, M. S. and A. K. T.; resources, M. S., A. K. T. and N. A.; data curation, M. S. and A. K. T; writing – original draft preparation, M. S., A. K. T. and N. A.; writing – review and editing, M. S., A. K. T. and N. A.; visualization, M. S. and A. K. T; supervision, M. S. and N. A.; project administration, M. S., A. K. T. and N. A.; funding acquisition – not applicable All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

Author Contributions: Conceptualization, C. Y.; methodology, C. Y.; formal analysis, C. Y.; investigation, C. Y.; resources, C. Y.; data curation, C. Y.; writing – original draft preparation, C. Y.;writing – review and editing, C. Y.; validation, Ş. Z. A.; Ş. Z. A.;visualization, C. Y.; supervision, C. Y. and Ş. Z. A.; project administration, C. Y.; funding acquisition – not applicable. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

References

1. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2020;395:709-733. https://doi.org/10.1016/S0140-6736(20)30045-3
2. Levin A, Stevens P, Bilous R, Coresh J, De Francisco ALM, De Jong PE, Griffith KE, Hemmelgarn BR, Iseki K, Lamb EJ, Levey AS, Riella MC, Shlipak MG, Wang H, White CT, Winearls CG. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl.* 2013;3(1):1-150. https://doi.org/10.1038/kisup.2012.73
3. Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, Saran R, Wang AYM, Yang, CW. Chronic kidney disease: global dimension and perspectives. *Lancet*. 2013;382(9888):260-272. https://doi.org/10.1016/S0140-6736(13)60687-X
4. Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. *Kidney Int Suppl.* 2022;12(1):7-11. https://doi.org/10.1016/j.kisu.2021.11.003

5. Mitema D, Jaar BG. How can we improve the quality of life of dialysis patients? *Semin Dial.* 2016;29(2):93-102. https://doi.org/10.1111/sdi.12467
6. Maaroufi A, Fafin C, Mougel S, Favre G, Seitz-Polski B, Jeribi A, Vido S, Dewisme C, Albano L, Esnault V, Moranne, O. Patients' preferences regarding choice of end-stage renal disease treatment options. *Am J Nephrol.* 2013;37(4):359-369. https://doi.org/10.1159/000348822
7. Wong G, Howard K, Chapman JR, Chadban S, Cross A, Tong A, Webster AC, Craig JC. Comparative survival and economic benefits of deceased donor kidney transplantation and dialysis in people with varying ages and co-morbidities. *PLoS One*. 2012;7(1):e29591. https://doi.org/10.1371/journal.pone.0029591
8. van der Tol A, Lameire N, Morton RL, Van Biesen W, Vanholder R. An international analysis of dialysis services reimbursement. *Clin J Am Soc Nephrol.* 2019;14(1):84-93. https://doi.org/10.2215/CJN.08150718
9. Centers for Disease Control and Prevention. Chronic Kidney Disease in the United States, 2023. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2023.
10. United States Renal Data System. 2022 USRDS Annual Data Report: Epidemiology of Kidney Disease in the United States. Bethesda, MD: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 2022.
11. Gerasimoula K, Lefkothea L, Maria L, Victoria A, Paraskevi T, Maria P. Quality of life in hemodialysis patients. *Mater Sociomed.* 2015;27(5):305-309. https://doi.org/10.5455/msm.2015.27.305-309
12. Brown L, Gardner G, Bonner A. A comparison of treatment options for management of end stage kidney disease in elderly patients: a systematic review. *JBI Evidence Synthesis.* 2014;12(7):374-404. https://doi.org/10.11124/jbisrir-2014-1152
13. Pramayanti NNT, Arnaya IK. The satisfaction analysis of kidney failure patients on the quality of hemodialysis services in the National Health Guarantee Era (JKN) (case study in Surya Husadha General Hospital Denpasar). *Int J Sci Basic Appl Res.* 2019;48(7):139-149. https://www.gssrr.org/index.php/JournalOfBasicAndApplied/article/view/10615 Accessed May 18, 2025.
14. Bahadori M, Raadabadi M, Jamebozorgi MH, Salesi M, Ravangard R. Measuring the quality of provided services for patients with chronic kidney disease. *Nephro-Urol Mon.* 2014;6(5):e21810. https://doi.org/10.5812/numonthly.21810
15. Bayhan B, Devebakan N. Patients' perceptions of service quality in the hemodialysis unit: A case study of Ege University Faculty of Medicine Hospital. [in Turkish] J Int Soc Res. 2015;8(41):1136-1146.
16. Bowen Bucheli A, Fosado Téllez O. Perception and expectations of patients from specialized renal health centers: SERVQUAL method and multivariate statistics. *Novasinergia.* 2020;3(2):30-39. https://doi.org/10.37135/ns.01.06.02
17. Palmer SC, de Berardis G, Craig JC, Tong A, Tonelli M, Pellegrini F, Ruospo M, Hegbrant J, Wollheim C, Celia E, Ferrari JN, Török M, Murgo M, Leal M, Bednarek-Skublewska A, Dulawa J, Strippoli GFM. Patient satisfaction with in-center haemodialysis care: An international survey. BMJ Open. 2014;4:e005020. https://doi.org/10.1136/bmjopen-2014-005020
18. Serrano-del Rosal R, Loriente-Arín N. The anatomy of patient satisfaction. *Salud Publica Mex.* 2008;50(2):162-172. https://doi.org/10.1590/s0036-36342008000200010
19. Rix EF, Barclay L, Stirling J, Tong A, Wilson S. The perspectives of Aboriginal patients and their health care providers on improving the quality of hemodialysis services: A qualitative study. *Hemodial Int.* 2015;19(1):80-89. https://doi.org/10.1111/hdi.12201
20. Grönroos CA service quality model and its marketing implications. *Eur J Mark.* 1984;18(4):36-44. https://doi.org/10.1108/EUM0000000004784
21. Parasuraman A, Zeithaml VA, Berry LL. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *J Retail.* 1988;64(1):12–37. https://doi.org/10.1016/S0022-4359(88)90021-8
22. Maghsoodi AI, Saghaei A, Hafezalkotob A. Service quality measurement model integrating an extended SERVQUAL model and a hybrid decision support system. *Eur Res Manag Bus Econ.* 2019;25:151-164. https://doi.org/10.1016/j.iedeen.2019.04.004
23. Yildiz O, Erdil S. Comparative measurement of service quality in the Turkish airline passenger transportation sector [in Turkish] Oneri. 2013;10(39):89-100. https://doi.org/10.14783/od.v10i39.1012000310
24. Marshall GN, Hays RD. The Patient Satisfaction Questionnaire Short-Form (PSQ-18). Santa Monica, CA: RAND; 1994. Available from: https://www.rand.org/health-care/surveys_tools/psq.html. Accessed May 18, 2025.
25. Kavalnienė R, Deksnyte A, Kasiulevičius V, Šapoka V, Aranauskas R, Aranauskas R. Patient satisfaction with primary healthcare services: Are there any links with patients' symptoms of anxiety and depression? *BMC Fam Pract.* 2018;19:90. https://doi.org/10.1186/s12875-018-0780-z
26. Lin DJ, Li YH, Pai JY, Sheu IC, Glen R, Chou MJ, Lee CY. Chronic kidney-disease screening service quality: questionnaire survey research evidence from Taichung City. *BMC Health Serv Res.* 2009;9:239. https://doi.org/10.1186/1472-6963-9-239
27. Hernández Meca M, Ochando García A, Mora Canales J, Lorenzo Martínez S, López Revuelta K. Satisfacción del paciente en una unidad de hemodiálisis: Objetivo de calidad asistencial en enfermería. [Spanish] *Rev Soc Esp Enferm Nefrol.* 2005;8(2):90-96.
28. Wasserfallen JB, Moinat M, Halabi G, Saudan P, Perneger T, Feldman HI, Martin PY, Wautersa JP. Satisfaction of patients on chronic haemodialysis and peritoneal dialysis. *Swiss Med Wkly.* 2006;136(13–14):210–217. https://doi.org/10.4414/smw.2006.11180
29. Yildirim A. The importance of patient satisfaction and health-related quality of life after renal transplantation. Transplant Proc. 2006;38(9):2831-2834. https://doi.org/10.1016/j.transproceed.2006.08.162
30. Meesala A, Paul J. Service quality, consumer satisfaction and loyalty in hospitals: Thinking for the future. *J Retail Consum Serv.* 2018;40:261-269. https://doi.org/10.1016/j.jretconser.2016.10.011
31. Al-Mhasnah A, Salleh F, Afthanorhan A, Ghazali P. The relationship between services quality and customer satisfaction among Jordanian healthcare sector. *Manag Sci Lett.* 2018;8(12):1413-1420. https://doi.org/10.5267/j.msl.2018.10.003
32. Hartanti LK, Antonio F. The effect of service quality dimensions on hemodialysis patient satisfaction in Indonesia. Indones J Health Adm. 2022;10(1):50-59. https://doi.org/10.20473/jaki.v10i1.2022.50-59
33. Elizar C, Indrawati R, Syah TYR. Service quality, customer satisfaction, customer trust, and customer loyalty in service of pediatric polyclinic over private H hospital of East Jakarta, Indonesia. *J Multidiscip Acad.* 2020;4(2):105–111. https://doi.org/10.20473/jaki.v10i1.2022.50-59

34. Kondasani RKR, Panda RK. Customer perceived service quality, satisfaction and loyalty in Indian private healthcare. *Int J Health Care Qual Assur.* 2015;28(5):452-467. https://doi.org/10.1108/IJHCQA-01-2015-0008
35. Yanuar Rahmat Syah T, Kusumo Wijoyo C. Service quality and customer satisfaction on WOM a private hospital in Indonesia. *J Medicoeticolegal Manaj Rumah Sakit.* 2024;10(1):22-34. https://doi.org/10.18196/jmmr.v10i1.10141
36. Zaim H, Bayyurt N, Zaim S. Service quality and determinants of customer satisfaction in hospitals: Turkish experience. *Int Bus Econ Res J (IBER).* 2010;9(5):51–58. https://doi.org/10.19030/iber.v9i5.8145
37. Aydemir E, Sahin Y. Evaluation of healthcare service quality factors using grey relational analysis in a dialysis center. *Grey Syst.* 2019;9(4):432-448. https://doi.org/10.1108/GS-01-2019-0001
38. Shaikh ZM. The impact of hospital accreditation on the patient's satisfaction of haemodialysis department services. *Int J Emerg Res Manag Technol.* 2017;6(8):384–392. https://doi.org/10.13140/RG.2.2.14818.02246

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17361

The Effect of Transtheoretical Model-Based Online Health Education on Nutrition and Physical Activity Change Processes in Overweight University Students: A Randomized Controlled Trial

Cansu Yaşar[1,2], Şeyma Zehra Altunkürek[3]

[1]*Department of Public Health Nursing, Nuh Naci Yazgan University, Kayseri, Turkey*
[2]*PhD Program, Erciyes University, Kayseri, Turkey*
[3]*Department of Public Health Nursing, Gülhane Health Sciences University, Ankara, Turkey*

Received: 2025-02-26.
Accepted: 2025-10-06.

J Clin Med Kaz 2025; 22(6): 28–36

Corresponding author:
Cansu Yaşar
E-mail: cyasar@nny.edu.tr.
ORCID: 0000-0003-1105-7136.

ABSTRACT

Background: Overweight and obesity have emerged as critical global public health issues, as they elevate the risk of non-communicable diseases. The university years constitute a pivotal phase characterized by the prevalence of detrimental food practices and inconsistent physical exercise, hence intensifying associated health concerns. This study seeks to assess the influence of an online health education program, grounded in the Transtheoretical Model, on the dietary and physical activity behavior modification processes of overweight university students.

Methods: This study was designed as a two-arm, parallel, randomized controlled clinical trial. The research was conducted at a foundation university in Turkey, involving 54 overweight students who satisfied the inclusion criteria and participated in training sessions once a week for one month. Participants were randomly assigned to intervention (n=27) and control (n=27) groups. The intervention group had online health education grounded in the Transtheoretical Model for four weeks, while the control group got no intervention. Measurements were taken before and after the test using the Nutrition Change Processes Scale, Exercise Change Processes Scale, and the Transtheoretical Model Stage of Change Form.

Results: The intervention group demonstrated a statistically significant increase in scores on the Nutrition Improvement Processes Scale and Exercise Change Processes Scale ($p<0.05$), while the control group exhibited no significant improvement. A larger percentage of students in the intervention group progressed to the action and maintenance stages compared to those in the control group.

Conclusions: The online health education based on the Transtheoretical Model significantly influenced dietary and physical activity behavior changes among overweight students. The results demonstrate that organized health education initiatives might enhance motivation and facilitate enduring behavioral change among university students. Registration: ClinicalTrials.gov Identifier NCT05482659

Keywords: Healthy Nutrition, Online Education, Overweight, Physical Activity, Transtheoretical Model.

Introduction

Within the scope of the Sustainable Development Goals, the UN has underlined how crucial it is to fight noncommunicable illnesses in order to ensure "Good Health and Well-Being" [1]. Consequently, the prevention of overweight and obesity—two significant risk factors for non-communicable diseases—is crucial for the protection and improvement of public health. As per the World Health Organization's (WHO) 2022 statistics, 43% of the global population is overweight, and over 16% is classified as obese [2]. In Turkey, 56% of people are overweight or obese overall, according to data from 2021 [3].

Overweight and obesity are not only diseases but also major risk factors for diabetes and cardiovascular disorders, various types of cancer, sleep disorders, gastrointestinal and respiratory diseases, musculoskeletal disorders, infertility, and various mental health conditions. They also lower quality of life and increase treatment costs [2, 4]. While obesity-related diseases affect all age groups and socioeconomic classes, studies report that risky health behaviors that lead to chronic diseases and mortality in adulthood often develop during university years [5-7]. According to studies, the following rates of overweight and obesity are common among college students: 48.6% in the US, 12–36% in Europe, 20.8% in Bangladesh, 2.9–14.3% in China, 20–30% in Malaysia, 31% in Thailand, 13–57.6% in Pakistan, 11–37.5% in India, 12.4–16.7% in Colombia, 31.6% in Mexico, 42% in Kuwait, 12.4% in Iran, and 11–40% in Turkey [7–10].

Factors influencing the development of overweight and obesity include age, gender, genetic predisposition, sociocultural factors (education, ethnicity, migration status), and lifestyle behaviors (physical activity and nutrition habits) [11-13]. During university years, environmental changes and increased stress levels trigger risky health behaviors, including unhealthy eating habits, smoking, alcohol consumption, and reduced physical activity [10, 14-16]. Research suggests that high-calorie food consumption among university students has increased, while their nutritional knowledge remains inadequate [17,18]. Promoting a nutritious diet and increasing physical exercise are essential components of adopting healthy lifestyle behaviors in the treatment of overweight and obesity [19]. T There are various models that have been developed to provide guidance in facilitating behavior change among individuals [20–22].

The Transtheoretical Model (TTM) was formulated by Prochaska and DiClemente in 1982 to augment individual motivation for behavioral modification [23]. In recent years, TTM has been recognized as an effective, practical, and applicable framework for encouraging behavioral change [24]. According to the model, individuals undergo specific stages of change when acquiring a new behavior. In TTM, an individual's stage of change is determined by their intention and motivation toward adopting a behavior. Each stage requires specific strategies and interventions, known as processes of change [25]. A review of previous studies indicates that TTM-based interventions have positively influenced the behavior change process among individuals with unhealthy eating behaviors and insufficient physical activity levels [24, 26]. However, most TTM-based nutrition and physical activity interventions conducted globally and in Turkey have focused on overweight and obese adults rather than university students [24, 27, 28]. Separate studies conducted in the United States, South Korea, and Malaysia have shown that TTM positively influenced university students' physical activity levels [29-31].

A literature review suggests that TTM-based nutrition and physical activity intervention studies targeting overweight university students remain limited. Since the university period is a critical stage for adopting healthy eating and adequate physical activity behaviors, previous research has recommended implementing intervention programs for university students. Accordingly, this study aims to assess the effectiveness of a health education program, structured based on the stages of change in the Transtheoretical Model, in facilitating nutrition and physical activity behavior change among overweight university students without any chronic diseases.

Methods

Design: This randomized controlled trial was conducted between June 2021 and May 2022 with a total of 54 students who met the inclusion criteria. An initial assessment (pre-test) was administered to overweight students at the commencement of the intervention, followed by a concluding assessment (post-test) after its conclusion to gather data. The randomized controlled study adhered to the Consolidated Standards of Reporting Trials criteria. This study was registered (registration number: NCT05482659) in the ClinicalTrials database.

Settings and Participants: The study population consisted of 2,520 students from a foundation university in Turkey. The sample size was determined using a repeated-measures two-way ANOVA in G*Power 3.1.9.2 to ensure statistical reliability. Due to the lack of similar prior studies, Cohen's (1988) recommended effect size (f = 0.25) was used, with a 0.05 Type I error rate and 0.95 statistical power. As a result, the minimum required sample size was calculated as 54 students (27 in the intervention group and 27 in the control group). The power analysis was conducted by a statistical expert [32]. The sample included participants from all academic years. Since this introduces variation in educational exposure and experience, academic level was considered a potential confounding variable. However, subgroup analysis revealed no statistically significant effect of academic level on outcomes. Participants living either in dormitories or family homes were included in the study, and statistical analysis revealed no significant effect of residence type on study outcomes.

Eligibility criteria for participation in the study:

• Individuals who voluntarily consented to take part in the study.

• Those with no chronic illness,

• BMI between 24.9 and 30,

• Those open to communication and collaboration,

• Those who own and have access to the internet, smartphone, tablet, or computer,

• Those who had not received a similar nutrition or physical activity intervention in the past 6 months.

Exclusion Criteria:

• BMI below 24.9 or above 30,

• Taking any medication,

Interventions: Students meeting the inclusion criteria were provided with an informed consent form via an online questionnaire, and those who voluntarily agreed to participate and fulfilled the eligibility requirements were enrolled in the study (n=54). The educational intervention was conducted in January 2022 and February 2022, after the completion of the semester examinations. This timing was selected to ensure that students were not under academic stress and could participate without distractions.

In the subsequent phase, a statistics expert assigned identification numbers to the students, which were then randomly allocated to the intervention and control groups using the Simple Randomization Method (Random Number Table). Participants were not informed about their group assignment (intervention or control) or their stage in the Transtheoretical Model (single-blind method). However, due to the nature of the intervention, the researcher was not blinded.

Students in the intervention and control groups were administered the pre-test survey through an online questionnaire (including the Demographic Information Form, Nutrition Change Processes Scale, Exercise Change Processes Scale, and Stage of Change Identification Form based on the TTM). After the pre-test, students were categorized into five groups according to the results of the Stage of Change Identification Form based on TTM:

• Precontemplation group: The person does not plan to modify their behavior within the next six months, has limited awareness, and may be resistant to change.

• Contemplation group: The individual has resolved to initiate the healthy activity within the next six months but requires further motivation and confidence enhancement.

• Preparation Group: Participants intend to adopt the healthy behavior within the next 30 days, actively seek information, and take steps to prepare for the change.

• Action Group: Participants have adopted the healthy behavior for less than six months but remain at risk of relapsing into previous habits.

• Maintenance Group: Participants have sustained the healthy behavior for more than six months and require ongoing support to prevent relapse.

Health Education Program: The educational program was developed by the researcher based on expert opinions from a Doctor of Public Health Nursing.

• The education was delivered online, once per week, to the intervention group.

• The content of the education program, structured according to change stages, is presented in Table 1.

• Each student received individualized, online training lasting 30-45 minutes on average.

• The intervention process lasted 30 days in total.

• The control group was not exposed to any experimental intervention.

• Final Test and Study Conclusion: Upon completion of the intervention process (30 days after its initiation), the post-test was administered to both groups (Nutrition Change Processes Scale, Exercise Change Processes Scale, and Stage of Change Identification Form based on TTM).

• Sharing of Educational Content: After the final test, students who requested access to the health education materials were provided with the content.

The study flowchart is shown in Figure 1.

Measurements: The study data were collected through a questionnaire developed based on a comprehensive literature review. The questionnaire consisted of the Demographic Information Form, Nutrition Change Processes Scale, Exercise Change Processes Scale, and Transtheoretical Model Stage of Change Identification Form. Permission was obtained for the use of the scales in the study.

The Demographic Information Form consists of 16 questions covering students' sociodemographic characteristics, height-weight values, nutrition, and physical activity habits.

The Nutrition Change Processes Scale (NCPS): Developed by Prochaska, Velicer, Rossi, Goldstein, Marcus, Rakowski, Fiore, Harlow, Redding, and Rosenbloom (1987), this scale assesses how experiences influence dietary habits. The Turkish adaptation, validated by Menekli and Fadıloğlu (2012). The scale consists of 12 subdimensions including 48 items evaluated on a 5-point Likert scale, with total scores varying from 48 to

Table 1 Stages of Nutrition and Exercise Change, Change Processes, and Training Content Provided According to the Transtheoretical Model

Stage of Change According to TTM	Change Processes According to TTM	Objective of Training	Content of Training	Method
Precontemplation	Emotional Arousal, Awareness	Increasing awareness of healthy eating and adequate physical activity	What is obesity/overweight, physical activity, and healthy eating? Determining the positive effects of healthy eating and physical activity on health. Healthy lifestyle behaviors	Q&A, Online Lesson
Contemplation	Awareness, Self-Reevaluation, Counter Conditioning	Motivating individuals for healthy eating and physical activity and increasing self-confidence	What is obesity/overweight, physical activity, and healthy eating? Barriers to healthy lifestyle behaviors, healthy eating, and physical activity. Healthy me? Unhealthy me?	Q&A, Online Lesson
Preparation	Self-Agreement, Self-Reevaluation, Environmental Reevaluation, Counter Conditioning	Preparing a plan for healthy eating and physical activity	What is obesity/overweight, physical activity, and healthy eating? Evaluating individuals' conditions related to healthy eating and physical activity and forming a plan	Q&A, Online Lesson
Action	Self-Agreement, Reinforcement (Rewarding), Stimulus Control	Starting a healthy eating and physical activity plan and revising the implementation plan	What is obesity/overweight, physical activity, and healthy eating? Evaluating individuals' conditions and starting a healthy eating and physical activity plan, environmental changes, and the importance of self-reward	Q&A, Online Lesson
Maintenance	Self-Agreement, Reinforcement (Rewarding), Stimulus Control	Preventing relapse into poor nutrition and inactivity	Making a contract, importance of not quitting exercise, stimulus control, evaluating individuals' healthy eating and physical activity plans, importance of self-rewarding	Q&A, Online Lesson

Q&A – Question and Answer.

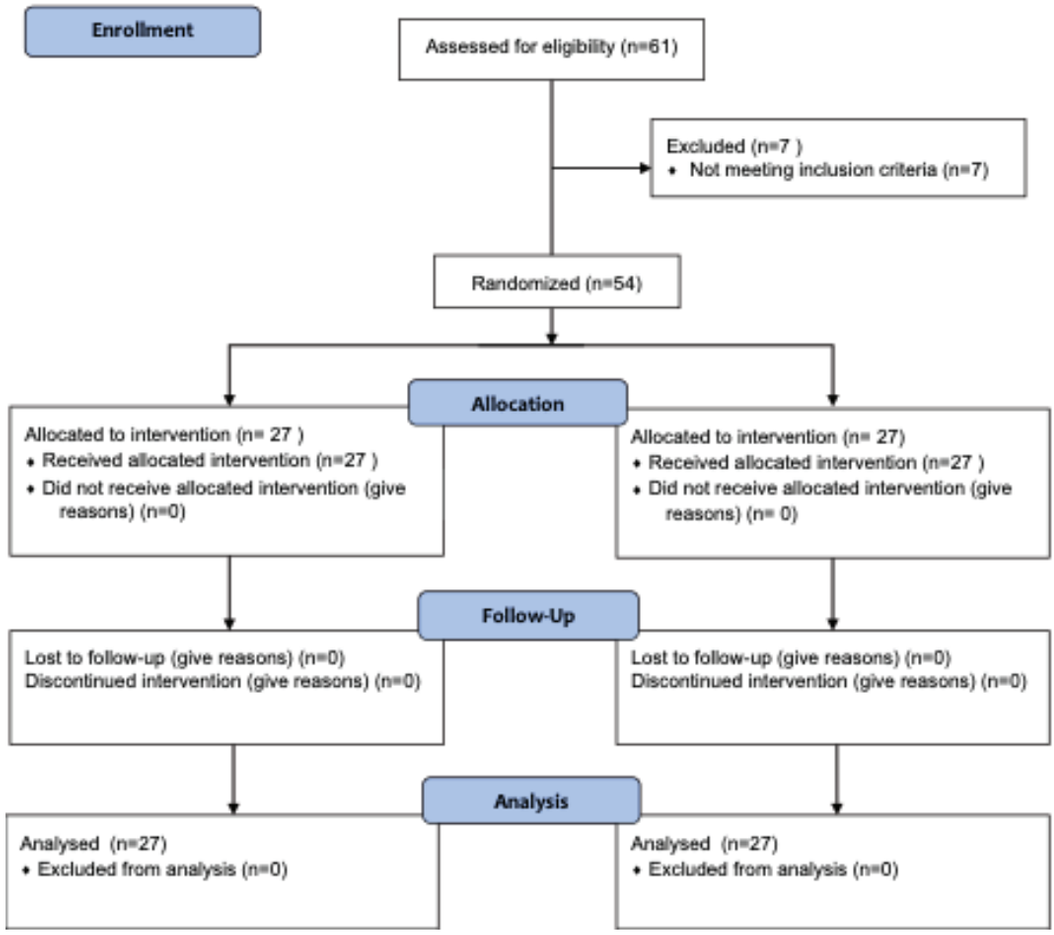


Figure 1 – Research Flowchart According to CONSORT 2010 Flow Diagram

240. In this study, its reliability was found to be Cronbach's Alpha = 0.93 [33].

The Exercise Change Processes Scale (ECPS): Created by Marcus, Selby, Niaura, and Rossi (1992), this scale evaluates the impact of experiences on exercise habits. Ay and Temel (2015) conducted the Turkish adaptation. The measure has two primary processes (cognitive and behavioral) spanning ten subdimensions, featuring 28 items evaluated on a 5-point Likert scale [34]. Scores range from 28 to 140, with higher scores indicating a greater likelihood of behavioral change. In this study, its reliability was Cronbach's Alpha = 0.932.

The Transtheoretical Model Stage of Change Identification Form was developed by the researcher based on a literature review to determine the stage of change for overweight individuals according to the Transtheoretical Model. Based on the results of this form, overweight students were classified into five groups: precontemplation, contemplation, preparation, action, and maintenance. Participants were asked, "Are you currently making or planning to make any changes to your nutrition and physical activity behaviors?" Their responses were classified into the following categories:

• "I am disinterested and have no intention of altering my dietary practices or exercise regimen in the forthcoming six months." (Precontemplation);

• "I am not interested at the moment, but I plan to change my nutrition habits and physical activity level within the next six months." (Contemplation);

• "I intend to modify my nutrition habits and physical activity level within the next 30 days." (Preparation).

• "Yes, I have already changed my nutrition habits and physical activity level, but it has been less than six months." (Action)

• "Yes, I have maintained my nutrition and physical activity behaviors for more than six months." (Maintenance)

Statistical analysis: The study data were examined utilizing SPSS version 22.0. Descriptive statistics, including mean, standard deviation, median, minimum, and maximum, were employed for continuous data, whilst categorical variables were represented as frequencies and percentages. The Kolmogorov-Smirnov test was employed to evaluate normality, affirming a normal distribution for the Nutrition and Exercise Change Processes Scales. Consequently, parametric tests were utilized: the t-test for independent group comparisons, the paired t-test for repeated measures, and the Chi-square test for categorical comparisons. The threshold for statistical significance was established at $p < 0.05$.

Results

At the commencement of the intervention, the chi-square test was employed to examine the sociodemographic features of the participants, indicating no statistically significant difference between the groups (Table 2, see the next page). Table 2 indicates that a predominant proportion of students in both the experimental (77%) and control (70.4%) groups were female. Furthermore, over fifty percent of the students (66.7%) reported that their income corresponded with their expenses.

Before the intervention, BMI and age values were analyzed, showing no significant differences between the experimental

Table 2 Distribution of Group Characteristics at the Beginning of the Intervention (n=54)

Characteristics	Variables	Intervention (n=27)		Control (n=27)		Statistics	
		n	%	n	%	χ2	p
Gender	Female	21	77.8	19	70.4	0,386	0,379
	Male	6	22.2	8	29.6		
Economic Status	Income > Expenses	9	33.3	8	29.6	1,059	0,589
	Income = Expenses	18	66.7	18	66.7		
	Income < Expenses	0	0.0	1	3.7		
Smoking Status	Smoker	10	37.0	11	40.7	0,078	0,500
	Non-smoker	17	63.0	16	59.3		
Alcohol Consumption	Yes	6	22.2	3	11.1	1,200	0,234
	No	21	77.8	24	88.9		
Weight Loss Status Before the Study	Yes	13	48.1	7	25.9	2,859	0,079
	No	14	51.9	20	74.1		
Body Satisfaction Status	Yes	12	44.4	15	55.6	0,667	0,293
	No	15	55.6	12	44.4		
Perceived Health Status	Yes	14	51.9	15	55.6	0,074	0,500
	No	13	48.1	12	44.4		
Number of Meals Consumed Per Day	1 meal	1	3.7	0	0.0	5,255	0,262
	2 meals	10	37.0	7	25.9		
	3 meals	11	40.7	16	59.3		
	4 meals	1	3.7	3	11.1		
	5 meals	4	14.8	1	3.7		
Average Daily Sleep Duration	6-8 hours	18	66.7	18	66.7	0,000	1,000
	8-10 hours	8	29.6	8	29.6		
	10-12 hours	1	3.7	1	3.7		
Frequency of Exercise Per Week	Never exercise	5	18.5	11	40.7	7,707	0,052
	Once a week	5	18.5	8	29.6		
	Twice a week	4	14.8	4	14.8		
	Three or more times a week	13	48.1	4	14.8		
Perceived Adequacy of Physical Activity Level	Yes	12	44.4	8	29.6	1,271	0,199
	No	15	55.6	19	70.4		

χ2 – Chi-Square Analysis, % – Percent, n – Number.

Table 3 Comparison of Students' BMI and Age Averages Between Groups (n=54)

Variables	Intervention (n=27)	Control (n=27)	Statistics	
	M±SD	M±SD	t	p
BKİ	27,7±1,4	27,4±1,6	0,824	0,520
Age	21,0±3,7	20,8±2,2	0,266	0,791

M – Mean, n – Number, SD – Standard Deviation, t – Independent Samples t-test

(BMI: 27.7±1.4, age: 21.0±3.7) and control groups (BMI: 27.4±1.6, age: 20.8±2.2) (p > 0.05, Table 3, see the next page).

Nutritional and Exercise Change Stages

Before the intervention, no significant difference was found between the experimental and control groups regarding nutritional and exercise change stages based on the TTM (p > 0.05, Table 3). In the experimental group, most students were in the preparation stage (n=19), while in the control group, the majority were also in preparation (n=20). After the intervention, a significant difference emerged (p = 0.001), with 20 students in

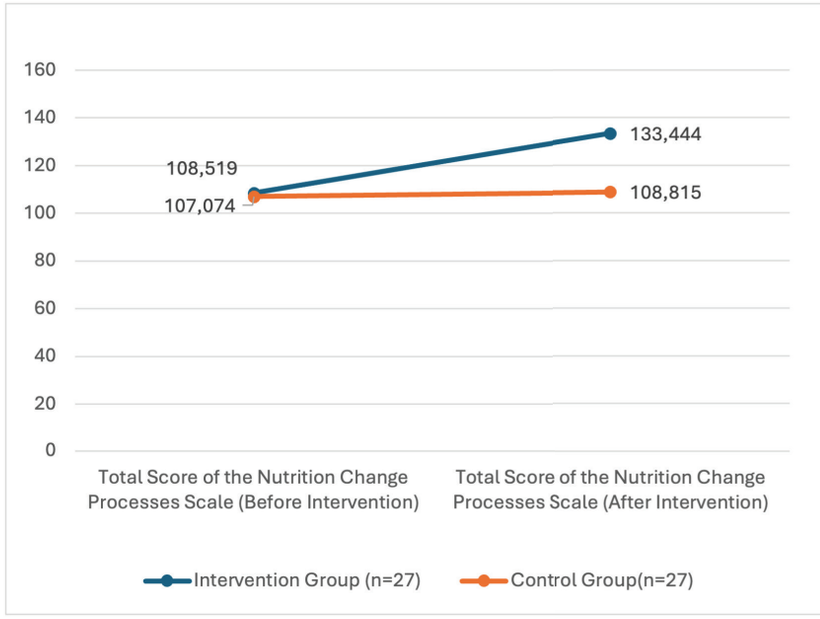


Figure 2 – Comparison of the Mean Scores of the Nutrition Change Processes Scale Between Groups Before and After the Intervention

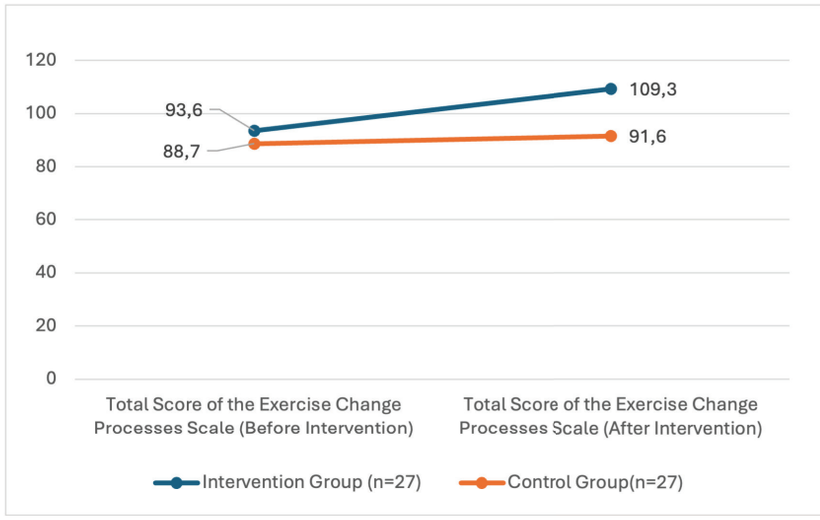


Figure 3 – Comparison of the Mean Scores of the Exercise Change Processes Scale Between Groups Before and After the Intervention

the experimental group progressing to the action stage, whereas the control group showed minimal change (Table 3).

Students' Nutritional Change Processes

Figure 2 presents the within-group comparison of the NCPS total and subscale mean scores before and after the intervention. In the experimental group, the mean score demonstrated a statistically significant increase from 108.5±30.1 to 133.4±36.8 ($t=3.821$, $p<0.001$). In the control group, the comparison of scores before and after the intervention revealed no statistically significant increase (107.0±27.4 to 108.8±31, $t=-0.801$, $p>0.05$).

Students' Exercise Change Processes

Figure 3 shows the within-group comparison of ECPS total and subscale mean scores pre- and post-intervention. In the experimental group, the mean score increased significantly from 93.6±17.3 to 109.3±17.8 ($t=-5.662$, $p=0.001$). In the control group, the comparison of scores before and after the intervention revealed no statistically significant increase (88.7±25.2 to 91.6±25.6, $t=-2.016$, $p=0.054$).

Discussion

In this study, the TTM was used as a guide to help individuals adopt healthy eating and regular physical activity behaviors. According to TTM, behavior change occurs through specific stages, and the stage an individual is in reflects their intention and motivation toward change [35]. In our study, when the pre- and post-intervention nutritional and exercise change stages of students in the experimental group were compared,

Table 4 Comparison of Nutrition and Exercise Change Stages Between Groups (n=54)

Stages of Change	Nutrition and Exercise	Intervention (n=27)		Control (n=27)		Statistics
		n	%	n	%	
Before Intervention	Precontemplation	0	%0	1	%3,7	χ2=3,026 p=0,554
	Contemplation	4	%14,8	1	%3,7	
	Preparation	19	%70,4	20	%74,1	
	Action	2	%7,4	3	%11,1	
	Maintenance	2	%7,4	2	%7,4	
After Intervention	Precontemplation	0	%0	1	%3,7	χ2=28,836 **p=0,001***
	Contemplation	0	%0	4	%14,8	
	Preparation	4	%14,8	18	%66,7	
	Action	20	%74,1	2	%7,4	
	Maintenance	3	%11,1	2	%7,4	

χ2 – Chi-Square Analysis, % – Percent, n – Number.

it was found that before the intervention, 14.8% of students were in the contemplation stage and 7.4% in the action stage, whereas after the intervention, these rates changed to 0% and 74.1%, respectively. Additionally, the fact that the education program lasted only one month suggests that TTM-based health education was effective in accelerating students' transition to healthy lifestyle behaviors in a short period.

Recent research indicates that TTM is an effective approach in promoting healthy eating and physical activity habits. Gereklioglu et al. (2024) found that a TTM-based intervention significantly increased Exercise Change Process Scale (ECPS) scores in their study on adults. Similarly, Abdoli et al. (2025) reported that TTM-based interventions were effective in increasing physical activity levels among female university students. These findings support the results of our study [26, 27, 36-38]. TTM is considered an easily applicable and effective guide for helping students develop healthy eating and exercise habits. One of the most significant advantages of the model is that it evaluates the change process in stages and includes interventions specific to each stage. For instance, while strategies to increase awareness are applied in the contemplation stage, motivational support techniques are used in the action stage. This flexible structure makes TTM more applicable than other behavior change theories.

However, for TTM-based education to be effective, individuals must be adequately supported according to their change processes. If an individual does not receive the necessary psychosocial and environmental support, behavior change may not be sustainable. Particularly due to a lack of motivation or external factors, individuals may revert to their previous habits. Therefore, identifying the stage individuals are in and applying appropriate education strategies tailored to each stage is crucial.

In this study, TTM-based health education was provided online and individually, which can be considered a factor increasing the feasibility of the intervention. Online health education has gained importance, especially in the post-pandemic period, due to its potential to reach large audiences, flexible scheduling, and individual follow-up advantages [39]. However, some studies have also highlighted the disadvantages of online education. For example, the possibility of participants losing motivation, lack of direct interaction, and varying levels of commitment among students are among the limitations of this method [40].

Conclusion

The study population was limited to students at a private university in Turkey. Therefore, the results can only be generalized to this specific group and cannot be applied to the entire population. It is recommended to use the Transtheoretical Model (TTM)-based health education in behavior change interventions. We suggest conducting further studies with larger sample sizes and longer durations to explore this topic in more depth. Since students from all academic years were included, differences in educational exposure could not be fully controlled and are therefore noted as a potential limitation.

It may be beneficial to develop practical strategies based on participants' economic, cultural, and social environments to support sustainable behavior change. The intervention in this study lasted one month; thus, longer follow-up periods (at least 3-6 months) are recommended to assess the sustainability of behavior changes. Additionally, studies comparing online education with face-to-face education can provide a better understanding of the effectiveness of this method.

Author Contributions: Conceptualization, C. Y.; methodology, C. Y.; formal analysis, C. Y.; investigation, C. Y.; resources, C. Y.; data curation, C. Y.; writing – original draft

preparation, C. Y.;writing – review and editing, C. Y.; validation, Ş. Z. A.; Ş. Z. A.;visualization, C. Y.; supervision, C. Y. and Ş. Z. A.; project administration, C. Y.; funding acquisition – not applicable. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: We sincerely thank all the students who voluntarily participated in this study. Their time, interest, and contributions were invaluable in successfully completing this research. Their motivation and determination to adopt healthy lifestyle behaviors played a crucial role in this study. We extend our gratitude to each and every one of them for their support.

Funding: None.

Data availability statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. United Nations Turkey. Good Health and Well-Being. Sustainable Development Goal 3. n.d. Accessed: https://www.globalgoals.org/goals/3-good-health-and-well-being/
2. WHO, Obesity and overweight. 2024. Accessed: https://www.clinmedkaz.org/home/instructions-for-authors
3. Organisation for Economic Co-operation and Development. Obesity, diet and physical activity. Retrieved: 26.02.2025, from https://www.oecd.org/en/topics/obesity-diet-and-physical-activity.html
4. Lehnert T, Sonntag D, Konnopka A, Riedel-Heller S, König HH. Economic costs of overweight and obesity. *Best Pract Res Clin Endocrinol Metab.* 2013;27(2):105-115. Epub 2013 Feb 28. PMID: 23731873. https://doi.org/10.1016/j.beem.2013.01.002
5. Andargie M, Gebremariam K, Hailu T, Addisu A, Zereabruk K. Magnitude of overweight and obesity and associated factors among public and private secondary school adolescent students in Mekelle City, Tigray Region, Ethiopia, 2019: Comparative cross-sectional study. *Diabetes Metab Syndr Obes*. 2021;14:901-915. PMID: 33688225; PMCID: PMC7936680 https://doi.org/10.2147/DMSO.S262480
6. Başçı, A.B., Emiroğlu, O.N. & Kalanlar, B. The effect of a nutrition program based on the Health Behavior Interaction Model on primary school students' nutritional attitudes and behaviors. *J Public Health (Berl.)* 2023. https://doi.org/10.1007/s10389-023-02101-1
7. Öcalan D, Ceylantekin Y, Kunduracılar Z, Doğan T. Investigation of the relationship between body mass index, burnout level, and well-being in university students. [In Turkish]. *Turkish Journal of Diabetes and Obesity.* 2020;4(3):270-278. https://doi.org/10.25048/tudod.685181
8. Özkan, İ, Adıbelli, D, İlaslan, E, Taylan, S. Relationship between obesity awareness and body mass index in university students. [In Turkish]. *Acıbadem University Journal of Health Sciences.* 2020;11(1):0-0. https://doi.org/10.31067/0.2020.249
9. Peltzer K, Pengpid S, Samuels TA, Özcan NK, Mantilla C, Rahamefy OH, Wong ML, Gasparishvili A. Prevalence of Overweight/Obesity and Its Associated Factors among University Students from 22 Countries. *International Journal of Environmental Research and Public Health.* 2014; 11(7):7425-7441. https://doi.org/10.3390/ijerph110707425
10. Arslan SA, Daşkapan A, Çakır B. Specification of nutritional and physical activity habits of university students [In Turkish]. *TAF Preventive Medicine Bulletin.* 2016;15(3):171-180. https://doi.org/10.5455/pmb.1-1436432564
11. Yang SJ, Choi HY, Chee YK, Kim JA. Prevalence and correlates of obesity and overweight among Asian immigrant women in Korea. *Asia Pac J Public Health.* 2012;24(4):620-630. https://doi.org/10.1177/1010539511406710
12. Park JH, Moon JH, Kim HJ, Kong MH, Oh YH. Sedentary lifestyle: Overview of updated evidence of potential health risks. *Korean J Fam Med.* 2020;41(6):365-373. https://doi.org/10.4082/KJFM.20.0165
13. Barlow P, Reeves A, McKee M, Galea G, Stuckler D. Unhealthy diets, obesity and time discounting: A systematic literature review and network analysis. *Obesity Reviews*. 2016;17(9):810-819. https://doi.org/10.1111/obr.12431
14. Lisha NE, Sussman S. Relationship of high school and college sports participation with alcohol, tobacco, and illicit drug use: A review. *Addictive Behaviors.* 2009;35:399-407. https://doi.org/10.1016/j.addbeh.2009.12.032
15. Sogari G, Velez-Argumedo C, Gómez MI, Mora C. College students and eating habits: A study using an ecological model for healthy behavior. *Nutrients*. https://doi.org/10.3390/nu10121823
16. Güllü M, Küçükkömürler S. An ınvestigation of carbohydrate consumption and BMI values of university students [In Turkish]. *Journal of Turkish Studies.* 2020;15(3):1889-1904. https://doi.org/10.29228/turkishstudies.43012
17. Şanlıer N, Konaklıoğlu E, Güçer E. Relationship between nutrition knowledge, habits, and behaviors of young people and body mass index [In Turkish]. *Gazi Faculty of Education Journal.* 2009;29(2):333-352.
18. Turganbaeva N, Isakova D, Tayfun A. A study on eating behaviors of university students. [In Turkish]. *Journal of Tourism and Gastronomy Studies*. 2021;9(1):366-376. https://doi.org/10.21325/jotags.2021.79
19. Li B, Gao S, Bao W, Li M. Effectiveness of lifestyle interventions for treatment of overweight/obesity among children in China: A systematic review and meta-analysis. *Front Endocrinol (Lausanne)*. 2022 Oct 31;13:972954.PMID: 36387871; PMCID: PMC9659900. https://doi.org/10.3389/fendo.2022.972954
20. Wright C, Barnett A, Campbell KL, Kelly JT, Hamilton K. Behaviour change theories and techniques used to inform nutrition interventions for adults undergoing bariatric surgery: A systematic review. *Nutr Diet.* 2022 Feb;79(1):110-128. doi: 10.1111/1747-0080.12728. Epub 2022 Feb 8. PMID: 35137502; PMCID: PMC9304301. https://doi.org/10.1111/1747-0080.12728
21. Ersin F, Bahar Z. The impact of health promotion models on early breast cancer detection behaviors: A literature review. [In Turkish]. *Dokuz Eylül University Nursing School Electronic Journal.* 2012;5(1):28-38.

22. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *American Journal of Health Promotion.* 1997;12(1):38-48. https://doi.org/10.4278/0890-1171-12.1.38
23. Prochaska JO, Di Clemente CC. Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy*. 1982;19(3):276-288. https://doi.org/10.1037/h0088437
24. Mastellos N, Gunn LH, Felix LM, Car J, Majeed A. Transtheoretical model stages of change for dietary and physical exercise modification in weight loss management for overweight and obese adults. *Cochrane Database of Systematic Reviews*. 2014;2014(2). https://doi.org/10.1002/14651858.CD008066.pub3
25. Glanz K, Rimer BK, Viswanath K. Health Behavior and Health Education. Theory, Research and Practice. 4th ed. San Francisco: *Jossey Bass AWiley Imprint;* 2008.
26. Hashemzadeh M, Rahimi A, Zare-Farashbandi F, Alavi-Naeini A, Daei A. Transtheoretical model of health behavioral change: A systematic review. *Iran J Nurs Midwifery Res.* 2019;24(2):83-90. https://doi.org/10.4103/ijnmr.IJNMR_94_17
27. Menekli T, Fadıloğlu Ç, Şenuzun Aykar F. Behavior Change in Obese Individuals: The Transtheoretical Model Approach. [In Turkish]. *İnönü University Journal of Health Services Vocational School.* 2019;1:39-57. https://doi.org/10.33715/inonusaglik.545565
28. Küçük S, Türeyen A, Yıldırım Y. Investigation of theory and model use in higher education theses in nursing related to diabetes: Systematic review. [In Turkish]. *Turkiye Klinikleri Journal of Nursing Sciences.* 2023;15(1):275-290. https://doi.org/10.5336/nurses.2022-89751
29. Jeon DJ, Kim KJ, Heo M. Factors related to stages of exercise behavior change among university students based on the transtheoretical model. *J Phys Ther Sci.* 2014;26(12):1929-1932. https://doi.org/10.1589/jpts.26.1929
30. Han H, Pettee Gabriel K, Kohl HW. Application of the transtheoretical model to sedentary behaviors and its association with physical activity status. *PLoS One.* 2017. https://doi.org/10.1371/journal.pone.0176330
31. Liu KT, Kueh YC, Kim Y, Kuan G. Structural relationship in transtheoretical model of behavioural change constructs among students at a Malaysian university campus. *Asian Journal of Sport and Exercise Psychology.* 2021;1(2):83-88. https://doi.org/10.1016/j.ajsep.2021.09.002
32. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods.* 2007;39(2):175-191.
33. Menekli T, Fadıloğlu Ç. Validity and Reliability of the Nutrition Change Processes Scale[In Turkish]. Hacettepe University Faculty of Health Sciences Nursing Journal. 2015;19(2):1-21. Accessed: https://dergipark.org.tr/tr/download/article-file/88618
34. Ay T, Değirmenci B, Temel AB. Adaptation of Transtheoretical Model-Based Exercise Behavior Change Scales for Turkish Adults: A Validity and Reliability Study [In Turkish]. *Izmir University Medical Journal.* 2017;1:29-45.
35. Prochaska JO, Redding CA, Evers KE. The transtheoretical model and stages of change. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. 4th ed. San Francisco: *Jossey-Bass*; 2008. p. 97-121.
36. Gereklioglu C, Topal K, Velipasaoglu H, Aksoy H. The effectiveness of the trans-theoretical model in managing adult obese and overweight individuals. *Pak J Med Sci.* 2024;40(4):685-
37. Abdoli M, et al. Determinants of physical activity among female students based on the transtheoretical model. *Sci Rep*. 2025;15(1):1496. https://doi.org/10.1038/s41598-024-83055-6
38. Karintrakul S, Angkatavanich J. A randomized controlled trial of an individualized nutrition counseling program matched with a transtheoretical model for overweight and obese females in Thailand. *Nutr Res Pract.* 2017;11(4):319-326. https://doi.org/10.4162/nrp.2017.11.4.319
39. Plotnikoff RC, et al. Effectiveness of interventions targeting physical activity, nutrition and healthy weight for university and college students: A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity.* 2015;12(1):1-10. https://doi.org/10.1186/s12966-015-0203-7
40. Mohammed R, Al Al-Muslimawi A. Evaluating the impact of distance learning systems on educational outcomes during the COVID-19 pandemic. *International Journal of Humanities and Social Sciences.* 2024;6(6):2717-8293. http://dx.doi.org/10.47832/2717-8293.32.2

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17363

Evaluation of Respiratory Symptoms and Ventilatory Function amongst Food Vendors in Calabar, Nigeria: a Cross-sectional Comparative Study

Emmanuel Jimmy[1], Obeten Ekpo[2], Bassey Ekeng[3], Chidimma Ahaneku[1], Bernard Bassey[1], Simon Ereh[1], Thelma Bassey[1], Love Okafor[1], Agbo James[1], Eshan Henshaw[1,4], Chibueze Njoku[1], Etete Petters[5]

[1]Department of Internal Medicine, University of Calabar Teaching Hospital, Calabar, Nigeria
[2]Department of Internal Medicine, University of Uyo Teaching Hospital, Uyo, Nigeria
[3]Department of Medical Microbiology and Parasitology, University of Calabar Teaching Hospital, Calabar, Nigeria
[4]Department of Internal Medicine, University of Calabar, Calabar, Nigeria
[5]Department of Internal Medicine, University of Uyo Teaching Hospital, Uyo, Nigeria

Received: 2025-07-08.
Accepted: 2025-10-11.

J Clin Med Kaz 2025; 22(6): 37–43

Corresponding author:
Bassey Ekeng.
E-mail: basseyewaekeng@gmail.com.
ORCID: 0000-0002-0509-2480.

ABSTRACT

Background: Data from studies evaluating lung functions in individuals exposed to fumes generated during the preparation of roasted food items is lacking in the Nigerian setting. We aimed to evaluate the ventilatory functions of participants with direct exposure to smoke from roasted food and compared the findings with their counterparts (controls).

Methods: We conducted a cross-sectional comparative study involving food vendors (cases) of smoked food and workers (controls) operating in the same vicinity but not directly exposed to the smoke. A spirometer was used to measure the forced expiratory volume in one second (FEV1), forced vital capacity (FVC) and peak expiratory flow rate (PEFR). Thereafter, the percentage predicted FEV1, FVC, and FEV1/FVC indices were used to determine ventilatory impairment in both cases and controls.

Results: A total of 600 participants were enrolled, of which half (50.0%) were cases while the remaining 50.0% were in the control group. Respiratory symptoms such as coughing and wheezing, and the duration of presentation were higher in cases compared with the controls, with a statistically significant difference; p = 0.003, p = 0.001 and p = 0.011, respectively. In addition, impaired ventilatory functions were significantly higher in the cases compared with the controls (p <0.001). Similarly, a significant proportion of the food vendors with ≥ 5 years of exposure had obstructive and mixed patterns of lung defects and a lesser proportion with normal lung functions compared with those with < 5 years of exposure (p = 0.009). Likewise, the mean FEV1 (p <0.001), FVC (p <0.001), and PEFR (p = 0.005) were significantly higher in controls compared with cases.

Conclusion: The direct exposure of food vendors to smoke may impair lung function. Driving awareness of this occupational hazard is very crucial as it is a common source of livelihood in the Nigerian setting.

Keywords: Lung defects, Peak expiratory flow rate, biomass fuel, spirometer

Introduction

Food vendors dealing with the processing and sale of smoked food items are commonly encountered in the Nigerian setting. These edibles are readily available and affordable, thus making it a lucrative business. However, the preparation of these items presents a health risk to the individuals involved due to their chronic exposure to smoke from the charcoal/wood used in the process. In addition, the non-use of personal protective equipment like face masks, probably due to a lack of awareness, further increases the risks of developing impaired ventilation.

Fumes derived from the burning of oil are a combination of several toxic gases and carcinogens [1-3]. While wood smoke is an important source of phenols, nitrogen dioxide, sulphur dioxide and benzopyrenes, oil fumes generate aldehydes, polycyclic aromatic hydrocarbons and alkanoic compounds [4]. The respiratory system often suffers major consequences as it is the portal of entry for these pollutants [4]. This is affirmed in a study from southern Nigeria, which documented the occurrence of chronic obstructive pulmonary disease (COPD) among women exposed to wood smoke in Bogota and Ile-Ife, respectively [2,5]. Besides the increased risk of respiratory symptoms associated with this exposure, a decline in lung function has been reported in some studies [6-9]. However, none of these studies were targeted at individuals involved in the processing of grilled foodstuffs [9-11], whereas data from studies conducted in other climes like Asia evidently show increased frequency of respiratory symptoms and a decline in lung functions due to exposure to toxic gases from grilled food substances [12,13]. Our study aimed to address this public health concern in our environment by assessing the pattern of respiratory symptoms and ventilatory function amongst food vendors in the Calabar metropolis.

Methods

Study setting

This study was carried out in Calabar, Cross River State, Nigeria. Calabar is the capital of Cross River State, in South-South Nigeria and has an area of 406 square kilometres. It has a population of 371,022 according to the 2006 census [14]. Calabar is divided into Calabar Municipal and Calabar South local government areas (LGAs). Calabar Municipal lies between latitude 04° 15′ and 5° N and longitude 8° 25′ E. It is bounded in the North by Odukpani LGA, in the North-East by the Great Kwa River and in the south, by the Calabar River and Calabar South LGA.

Study population

The study population comprised food vendors located in Calabar, regardless of their local government. Each made use of tables to display their food items (fish, meat, plantain, yam) and used a make-shift grill over a metal container containing chunks of charcoal and wood pellets to provide heat for food processing, Figure 1.

Figure 1 – A grill overlying a container with chunks of charcoal and wood pellets to provide heat for food processing

Study design

This was a cross-sectional comparative study involving vendors (cases) of smoked food and workers (controls) operating in the same vicinity but not exposed to smoke.

Inclusion and exclusion criteria

The inclusion criteria for both cases and controls were 1. Consenting food vendors, 18 years and above. 2. Food vendors who have been in the business for at least 1 year. Likewise, exclusion criteria for both cases and controls included a history of chronic lung diseases such as COPD, smoking and asthma before the study. 2. Participants below 18 years.

Study instruments

1. CONTEC SP10 Spirometer (Contec, China, 2016), 2. Peak flow meter, mini-Wright (England), 3. The American Thoracic Society and National Heart and Lung Institute-Division of Lung Disease Respiratory Questionnaire (ATS-DLD-78-A)-interviewer administered [15], 4. Standard weighing scale (HANA weighing scale, Germany), 5. Tape measure (Inelastic), and 6. A Pulse Oximeter (Excelvan fingertip pulse Oximeter, Bio-Tek, China).

Sampling method

A total of 300 food vendors were sampled using the proportionate sampling method. One hundred and fifty food vendors were selected from Calabar South and Calabar municipality respectively. Calabar South has 11 wards, so the total number of food vendors sampled from each ward was 14 (150 food vendors divided by 11). Each ward has an average of 6 to 8 streets, so 2 to 3 food vendors were selected randomly from each street. A total of 14 food vendors per ward were selected. Calabar Municipality has 10 wards, so the total number of food vendors sampled from each ward was 15, with 2 to 3 food vendors selected randomly from each street in the ward. Each subject was recruited after a written consent was obtained. The biodata collected for the cases included the following parameters: age, height, weight, gender, and level of education.

As in the cases, the controls were chosen using the proportionate sampling method. The total number of controls was 300 (150 controls from Calabar South and 150 controls from Calabar Municipal). Each control was recruited after written consent was obtained. The biodata collected for the controls included the following parameters: age, height, weight, gender, and level of education.

Data collection

An ATS-DLD-78-A [15] was used to collect information regarding the subjects' respiratory symptoms after written consent had been obtained from them. The questionnaire was used to collect data on the socio-demographic characteristics of the participants (age, sex, education levels), duration of exposure, and presence of respiratory symptoms such as cough, wheezing and dyspnoea. In addition, a history of cigarette smoking, prior history of asthma, chronic bronchitis and use of personal protective equipment was also obtained. Data collection was primarily conducted by the lead author and assisted by two research assistants who were trained on how to collect data from the participants using the questionnaires. The local English language (Pidgin) and Standard English language were used to get the data from the participants.

Procedure

The technique involved a maximum inspiration followed by a forced expiration for as long as possible into the Spirometer. The procedure was performed by the researcher and thoroughly explained to the participants, followed by a demonstration before engaging the participants. The peak flow rate was measured using a Mini-Wright peak flow meter (England). The

participants were asked to inhale to full lung capacity, then seal their lips tightly around the mouthpiece with a nose clip occluding the nostrils and then exhale fast and forcefully in one blow into the mouthpiece. The plunger on the peak flow meter was noted and recorded. This was done either in groups or individually. Testing on each subject was done for a maximum of 6 times. The best 3 readings were recorded and the best reading of the 3 was used as the subject's best. About 15-30 minutes was allowed for resting after each session (six trials) for subjects unable to provide acceptable readings after each session and for subjects unable to perform the test adequately, it was repeated on a different day. Calibration was done daily before commencing testing on subjects. Before this, the patient's biodata (name, sex, age, height, and weight) was recorded in the Spirometer. The Pulse oximeter was used to measure the subject's oxygen saturation. The height (in meters) was measured with an inelastic tape whose calibration was checked with a Stadiometer, and the weight (in kilogram) was measured using a standard weighing scale (HANA weighing scale, Germany).

Evaluation of respiratory functions

A ventilatory function test was carried out using a Spirometer (CONTEC SP10, made in CHINA, 2016) by the researcher to evaluate for peak expiratory flow rate (PEFR), forced vital capacity (FVC), forced expiratory volume in one second (FEV1), and FEV1/FVC ratio. The FEV1 is the maximum volume of air in the first second of a forced expiration from a position of full inspiration. The FVC is the maximum volume of air exhaled with a maximally forced effort from a position of maximal inspiration. The PEFR is the maximum expiratory flow achieved from a maximum forced expiration starting without hesitation from a point of maximal lung inflation. The standard unit for PEFR is litres/minute. The ratio of the two figures (FEV1/FVC) tells us about the degree of airflow obstruction.

The prediction formulae for lung function parameters in females of South Eastern Nigeria [16], was used to calculate the predicted FEV1, FVC, and PEFR for the female subjects. The prediction equation for the various lung function indices in the female subjects were as follows: FVC = 0.145 + (1.390 x Height) – (0.0076 x Age) + (0.0089 x Weight), FEV1= 0.240 + (1.045 x Height) – (0.0055 x Age) + (0.0064 x Weight), PEFR= -38.80 + (210.83 x Height) + (1.650 x Age) + (0.252 x Weight).

The reference equation for the spirometric indices from a sample of the general adult population in Nigeria [17] was used to calculate the predicted FEV1, FVC and PEFR in males. The prediction equation for the various lung function indices in the male subjects was as follows: FEV1 = -0.834 + (-0.031) Age + 0.031 (Height),

FVC = -0.848 + (-0.034) A + 0.035 (Height), PEFR = -4.199 + (-0.054) A + 0.091 (Height).

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines was used to estimate the ventilatory impairment in the cases and controls [18]. The percentage predicted FEV1, FVC and FEV1/FVC indices were used to determine the ventilatory impairment in both the cases and controls. FEV1 less than 80% of predicted or FEV1/FVC ratio less than 0.7 of predicted indicated an obstructive pattern. A proportional reduction in the FEV1 and FVC with a normal or increased FEV1/FVC ratio was regarded as a restrictive pattern. A reduction in the FEV1 and FVC and a reduced FEV1/ FVC ratio was regarded as a mixed pattern.

The peak expiratory flow rate for the cases and controls was graded into 3. Subjects with predicted PEFR greater than 80% were grouped as Normal ventilatory function. Subjects with predicted PEFR between 50-80% were grouped as Mild ventilatory impairment. Participants with a predicted PEFR of less than 50% were grouped as Moderate ventilatory impairment and subjects with less than 30% predicted were grouped as severe ventilatory impairment.

Statistical analysis

Data collected was entered into an Excel spreadsheet and imported into Statistical Package for Social Sciences (SPSS) version 23.0 for cleaning and analysis. Descriptive statistics were presented in frequencies and percentages. The mean and standard deviations of continuous variables were analysed and compared using the student's t-test while the relationship between qualitative data was assessed using the Chi-square test or Fisher's Exact Test where the expected cell has a count less than 5. PEFR, FVC, FEV1, FEV1/FVC were measured and presented as mean and standard deviation. The association between duration of exposure with the ventilatory function changes among food vendors was carried out using the Student T-Test to assess for association between continuous variables. The level of significance was fixed at p-value <0.05.

Results

Socio-demographic characteristics of participants

A total of 600 individuals participated in the study. 50.0% (n=300) were cases (food vendors) while the remaining 50.0% (n=300) were in the control group. Amongst the cases, 50.0% (n=150) were males and amongst the controls, 61.7% (n=185) were males. Overall, males comprised 55.8% (n=335) of the participants compared with 44.2% (n=265) of females. The mean age in the control group was higher than that of cases, 30.8 ± 11.2 versus 27.4 ± 11.7, respectively. The predominant level of education across all participants was secondary school level of education (68.0%, n=408), Table 1.

Table 1 Socio-demographic characteristics of participants (N=600)

Variable	Cases n=300	Control n=300	Total N=600	Chi-square Test	p-value
Sex					
Male	150 (50.0)	185 (61.7)	335 (55.8)	8.279	0.057
Female	150 (50.0)	115 (38.3)	265 (44.2)		
Age group (years)					
15-24	164 (54.7)	100 (33.3)	264 (44.0)	33.406	0.062
25-34	70 (23.3)	116 (38.7)	186 (31.0)		
35-44	29 (9.7)	50 (16.7)	79 (13.2)		
45-54	26 (8.7)	21 (7.0)	47 (7.8)		
55-64	7 (2.3)	7 (2.3)	14 (2.3)		
>64	4 (1.3)	6 (2.0)	10 (1.7)		
Mean age ± SD	27.4 ± 11.7	30.8 ± 11.2	29.1 ±11.6	T-Test	0.073
Highest education					
None	4 (1.3)	1 (0.3)	5 (0.8)	30.767	<0.001*
Primary	54 (18.0)	60 (20.0)	114 (19.0)		
Secondary	226 (75.3)	182 (60.7)	408 (68.0)		
Tertiary	16 (5.3)	57 (19.0)	73 (12.2)		

**=statistically significant*

Comparison of the pattern and frequency of respiratory symptoms between cases and controls

Compared with the control group (40.7%, n=122), the proportion of participants who presented with cough was higher among the cases (52.7%, n=158) with a statistically significant difference (p=0.003).

Participants presenting with productive cough, a prolonged duration (≥ 3 weeks) of coughing, and wheezing were more amongst cases compared with controls and statistically significant with P values of 0.03, 0.011 and <0.001 respectively. Another clinical feature compared between the cases and controls was breathlessness: walking uphill, walking slowly, and a scenario where participants stopped breathing. In all, a significantly higher proportion of cases were affected compared with the controls, as shown in Table 2.

Table 2 Comparison of pattern and frequency of respiratory symptoms between cases and controls

Symptoms	Cases n=300	Control n=300	Total N=600	Chi-square Test	p-value
Cough					
Yes	158 (52.7)	122 (40.7)	280 (46.7)	9.398	0.003*
No	142 (47.3)	178 (59.3)	320 (53.4)		
Coughed for 3 Months					
Yes	41 (13.7)	22 (7.3)	63 (10.5)	6.402	0.011*
No	259 (86.3)	278 (92.7)	537 (89.5)		
Coughing with sputum for ≥ 3 weeks					
Yes	48 (16.0)	30 (10.0)	78 (13.0)	4.775	0.029*
No	252 (84.0)	270 (90.0)	522 (87.0)		
Wheezing					
Yes	75 (25.0)	38 (12.7)	113 (18.8)	14.926	<0.001*
No	225 (75.0)	262 (87.3)	487 (81.2)		
Breathless on walking uphill					
Yes	184 (61.3)	102 (34.0)	286 (47.7)	44.925	<0.001*
No	116 (38.7)	198 (66.0)	314 (52.3)		
Breathless on walking slowly					
Yes	63 (21.0)	43 (14.3)	106 (17.7)	4.583	0.032*
No	237 (79.0)	257 (85.7)	494 (82.3)		
You stop to Breath					
Yes	46 (15.3)	21 (7.0)	67 (11.2)	10.501	0.001*
No	254 (84.7)	279 (93.0)	533 (88.8)		

**=statistically significant*

Comparison of the pattern of ventilatory function between cases and controls

The proportion of cases (71.7%, n=215) with normal ventilatory function was less than that of controls (78.0%, n=235), and a higher frequency of mild ventilatory defects was observed amongst the cases (28.3%, n=85) compared with the controls (22.0%, n=66). However, the difference in the PEFR category between cases and controls was not statistically significant (p=0.074). In contrast, looking at lung defects, a higher proportion of participants amongst the controls (96%, n=288) were without lung defects compared with cases (71.7%, n=215) and vice versa with a statistically significant difference (p<0.001), Table 3.

Table 3 Comparison of pattern of ventilatory function test between cases and controls

Variable	Cases n=300	Control n=300	Total N=600	Chi-square Test	p-value
PEFR					
Normal	215 (71.7)	234 (78.0)	449 (74.8)	3.195	0.074
Mild defect	85 (28.3)	66 (22.0)	151 (25.2)		
Lung defects					
None	215 (71.7)	288 (96.0)	503 (83.8)	67.988	<0.001*
Obstructive	48 (16.0)	11 (3.7)	59 (9.8)		
Restrictive	20 (6.7)	1 (0.3)	21 (3.5)		
Mixed	17 (5.7)	0 (0.0)	17 (2.8)		

**=statistically significant, PEFR: Peak expiratory flow rate*

Association between the duration of exposure and changes in ventilatory function among food vendors

Among those who were exposed for less than 5 years (n = 142), a greater proportion had no lung defects (n = 109, 76.8%). The remainder had varied defects including obstructive (9.2%, n=13), restrictive (9.2%, n=13) and mixed (4.9%, n=4) lung defects. In contrast, those exposed for 5 years and above (n=158), had a relatively smaller proportion of participants (n=106, 67.1%) with no lung defects. Similarly, except for restrictive defects (4.4%, n=7), the frequency of obstructive (22.2%, n=35) and mixed (6.3%, n=10) defects were higher in cases with ≧5 years of exposure compared with cases with < 5 years of exposure, Table 4.

Table 4 Association between the duration of exposure with the ventilatory function changes among food vendors (n=300)

Duration of Exposure	None	Lung defects					
		Obstructive	Restrictive	Mixed	Total	Chi-square test	P-Value
<5 years	109 (76.8)	13 (9.2)	13 (9.2)	4 (4.9)	142 (100.0)	11.634	0.009*
≥5 years	106 (67.1)	35 (22.2)	7 (4.4)	10 (6.3)	158 (100.0)		

**=statistically significant*

The correlation between change in oxygen saturation (SPO_2) and ventilatory functions

The correlation between change in oxygen saturation (SPO_2) and ventilatory functions; FEV1, FVC and FEV1/FVC, was negative but however, not statistically significant (p>0.05). On the other hand, the correlation between oxygen saturation and PEFR was positive but not statistically significant (p=0.556), (Table 5).

Table 5 Association between the duration of exposure with the ventilatory function changes among food vendors (n=300)

SPO_2	PEFR(L/s)	FEV1(L/s)	FVC(L/s)	FEV1/FVC
Pearson correlation (r)	0.024	-0.002	-0.002	-0.018
P-value	556	954	952	655
N	600	600	600	600

PEFR: Peak expiratory flow rate, FEV1: forced expiratory volume in 1 second, FVC: Forced Vital Capacity, N=Total number of participants

Comparison of mean FEV1(L/s), FVC(L/s) and PEFR(L/s) between cases and control

The Mean PEFR was higher in the control group than in the cases (423.3 ± 85.2 versus 404.1 ± 82.7), and the difference was statistically significant (p=0.005). Likewise, the FEV1 and FVC were also higher among the controls than among the cases, and the differences were significant (p < 0.001), as shown in Table 6.

Table 6 Comparison of mean FEV1(L/s), FVC(L/s) and PEFR(L/s) between cases and control

Variables	Case	Control	Total	T-Test statistics	p-value
FEV1	2.7 ± 0.6	3.1 ± 0.6	2.9 ± 0.6	7.967	<0.001*
FVC1	3.4 ± 0.7	3.7 ± 0.7	3.6 ± 0.7	5.077	<0.001*
PEFR	404.1 ± 82.7	423.3 ± 85.2	413.7 ± 84.5	2.796	0.005*

**=statistically significant, PEFR: Peak expiratory flow rate, FEV1: forced expiratory volume in 1 second, FVC: Forced Vital Capacity*

Discussion

This study compared the respiratory symptoms and ventilatory impairment patterns amongst individuals directly exposed to solid biomass fuel and controls. Lung defects and chest symptoms were observed to be significantly common among the cases compared with the controls. Closely linked with this is the duration of exposure to solid biomass fuel, which showed a strong correlation with the frequency of symptoms and lung defects amongst the cases enrolled in this study. The most encountered defects and presenting symptoms amongst cases and controls were obstructive forms and coughing.

Looking at sociodemographic characteristics, females constituted 44.2% and males constituted 55.8% while about 75% of the total participants were below 35 years. This may be due to the stressful nature of the occupation, which requires muscular strength and endurance. A similar study by Adewole et al [19] from Northern Nigeria had males as the only participants, probably because the sale of Suya is a major occupation of the menfolk, whereas women in the north, for cultural and religious reasons, are not allowed into the occupation that exposes them to public glare. However, the predominance of females has been reported in some studies, which may be due to a variation in socio-cultural settings and perhaps the need to earn a living [20]. As per the level of education attained by participants in this study, the predominant level of education among both groups was the secondary level of education. This was different from an indoor study that had the majority of the participants in the primary level of education [21]. However, the nature of the occupation does not require any form of expertise; therefore, the level of education has no influence on the individuals involved in the business.

The clinical features observed amongst participants in this study was predominantly coughing, which is in keeping with findings from previous studies. Others were wheezing and dyspnoea, which were worse among the cases compared with the controls. A similar study done by Regalado et al had cough as the predominant feature among the cases [22]. Another study had chest tightness as the predominant symptom observed among the participants [19], and in yet another study done in the Niger Delta region, coughing was observed to be the predominant symptom among the participants [21]. From other climes, a Brazilian study reported an increased risk of chronic cough in adults exposed to biomass fuel [23], and from Brunei, cough was reported as the predominant respiratory symptom experienced in cooking vendors [13]. This corroborates the findings from our study, showing that exposure to smoke from solid biomass fuel is a risk factor for the development of respiratory symptoms in participants.

Regarding the impairment of ventilatory functions, besides the findings already narrated, the majority of the controls had normal PEFR when compared with the cases. The cases had a larger proportion of mild impairment in the PEFR when compared with the controls. Similarly, the mean PEFR, FEV1 and FVC were reduced in the cases compared with the controls. Earlier studies by Victor et al, Adewole et al, and Rothman et al revealed similar ventilatory impairment among workers exposed to smoke from biomass fuels, with obstructive ventilatory impairment as the predominant presentation [19,21,24]. Exposure to biomass fuel is therefore a major risk factor for the development of chronic obstructive ventilatory defects [9,10,22]. In addition, it is also associated with restrictive lung impairment (due to diffuse parenchymal fibrosis). This may explain the restrictive patterns observed in both groups. The mixed ventilatory pattern observed may be due to both airway broncho-constriction and diffuse parenchymal changes [25].

The duration of exposure was also a risk factor for increased respiratory symptoms and impaired ventilatory function. This may be due to a permanent structural change in the airways with subsequent persistent bronchoconstriction [26]. A larger proportion of cases exposed for more than 5 years had more ventilatory impairment when compared with cases exposed for less than 5 years. An obstructive ventilatory impairment was the predominant ventilatory defect observed. Earlier studies conducted in the Nigerian setting revealed similar findings [19,21]. Likewise, a study from Brazil reported an increased risk of developing chronic cough (OR 2.33, 95% CI 1.68-5.10), wheezing (OR 2.33, 95% CI 1.25-4.38) and dyspnoea (OR 2.59, 95% CI 1.32-5.09) in adults exposed to biomass smoke for more than 5 years [23]. The PEFR values were also observed to be worse in subjects exposed for more than 5 years in both cases and the controls. Thus, prolonged exposure to wood smoke may suggest a more severe impairment of lung function in affected individuals.

Lastly, on oxygen saturation, expectedly, the mean oxygen saturation was lower in the cases when compared with the controls. Oxygen saturation changes in an individual exposed to biomass fuel is usually significant in subjects with severe ventilatory impairment as mild and moderate impairment may still have normal oxygen saturation. The majority of the subjects in our study had a mild ventilatory impairment. This may have accounted for the normal oxygen saturation observed in our study.

Limitations

The concentration of the particulate matter in the air around where the food vendors operate was not measured to determine the effect of the different concentrations of the particulate matter on the respiratory symptoms and ventilatory function. It is possible that some areas had a higher concentration of particulate matter, and that may have influenced the findings we observed

Conclusion

Food vendors directly exposed to biomass fuel are at risk of developing impaired ventilatory function. Urgent steps should be taken to mitigate this through driving awareness and educating the public on the health hazards associated with this

occupational lifestyle and the adequate measures taken to prevent or minimize exposure. Ventilated grills can reduce the amount of exposure to smoke and subsequent lung impairment. In addition, food vendors should be encouraged to use protective devices to minimize the risk of respiratory diseases. Policies to encourage and enforce the use of protective devices should be encouraged among these at-risk groups of workers. Food vendors who have developed respiratory symptoms should be encouraged to go for further evaluation and treatment, and an annual spirometry evaluation and assessment of their respiratory status should be encouraged.

List of abbreviations

COPD: Chronic obstructive pulmonary disease

LGA: Local government areas

UCTH: University of Calabar Teaching Hospital, Calabar, Nigeria

ATS-DLD-78-A: American Thoracic Society and National Heart and Lung Institute-Division of Lung Disease Respiratory Questionnaire

PEFR: Peak expiratory flow rate

FVC: Forced vital capacity

FEV1: Forced expiratory volume in one second

GOLD: Global Initiative for Chronic Obstructive Lung Disease

SPSS: Statistical Package for Social Sciences

Author Contributions:

E.J.; Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing-original draft, Writing – review & editing.

O.E., B.B.; Methodology, Investigation, Visualization, Resources, Writing – review & editing,

B.E.: Methodology, Investigation, Formal analysis, Project administration, Visualization, Validation, Resources, Writing – review & editing,

C.A., S.E.: Data curation, Investigation, Visualization, Resources, Writing- review & editing,

T.B., L.O., A.J.; Investigation, visualization, Resources, Writing – review & editing, E.H.; Investigation, Methodology, Project administration, Resources, Visualization, Supervision, Writing-original draft, Writing – review & editing,

C.N., E.P.: Investigation, Methodology, Project administration, Resources, Validation, Visualization, Supervision, Writing – original draft, Writing – review & editing.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon resonable quest. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. Li S, Pan D, Wang G. Analysis of polycyclic aromatic hydrocarbons in cooking oil fumes. *Arch Environ Health*. 1994;49(2):119-22. https://doi.org/10.1080/00039896.1994.9937464
2. Erhabor GE, Kolawole OA, Adewole OO. Chronic obstructive pulmonary disease in OAUTHC Ile-Ife: a ten-year review of hospital mortality. *J Natl Med Assoc*. 2002 ;94(12):1071-6.
3. Tzanakis N, Kallergis K, Bouros DE, Samiou MF, Siafakas NM. Short-term effects of wood smoke exposure on the respiratory system among charcoal production workers. *Chest*. 2001;119(4):1260-5. https://doi.org/10.1378/chest.119.4.1260
4. Broaddus CV, Mason RJ, Ernst JD, King Jr TE, Lazarus SC, Murray JF, Nadel JA, Slutsky AS, Gotway MB. Murray and Nadel's textbook of respiratory medicine. 6th ed. *Elsevier;* 2016. 1-1849.e8
5. Dennis RJ, Maldonado D, Norman S, Baena E, Martinez G. Woodsmoke exposure and risk for obstructive airways disease among women. *Chest*. 1996;109(1):115-9. https://doi.org/10.1378/chest.109.1.115
6. O'Connor G, Neas L, Vaugh BJ, Kattan M, Mitchell H, Mastin P, Crain EF, Evans R, Gruchalla RS, Morgan W, Stout JW. Acute respiratory health effects of air pollution on asthmatic children in US inner cities. *Epidemiology*. 2004;15(4):S36-7. https://doi.org/10.1016/j.jaci.2008.02.020
7. Thiébaud HP, Knize MG, Kuzmicky PA, Hsieh DP, Felton JS. Airborne mutagens produced by frying beef, pork and a soy-based food. *Food Chem Toxicol.* 1995;33(10):821-8. https://doi.org/10.1016/0278-6915(95)00057-9
8. Svedahl S, Svendsen K, Qvenild T, Sjaastad AK, Hilt B. Short term exposure to cooking fumes and pulmonary function. *J Occup Med Toxicol.* 2009;4:9. https://doi.org/10.1186/1745-6673-4-9
9. Gordon SB, Bruce NG, Grigg J, Hibberd PL, Kurmi OP, Lam KB, Mortimer K, Asante KP, Balakrishnan K, Balmes J, Bar-Zeev N, Bates MN, Breysse PN, Buist S, Chen Z, Havens D, Jack D, Jindal S, Kan H, Mehta S, Moschovis P, Naeher L, Patel A, Perez-Padilla R, Pope D, Rylance J, Semple S, Martin WJ 2nd. Respiratory risks from household air pollution in low- and middle-income countries. *Lancet Respir Med.* 2014;2(10):823-60. https://doi.org/10.1016/S2213-2600(14)70168-7
10. Po JY, FitzGerald JM, Carlsten C. Respiratory disease associated with solid biomass fuel exposure in rural women and children: systematic review and meta-analysis. *Thorax*. 2011;66(3):232-9. https://doi.org/10.1136/thx.2010.147884
11. Naeher LP, Brauer M, Lipsett M, Zelikoff JT, Simpson CD, Koenig JQ, Smith KR. Woodsmoke health effects: a review. *Inhal Toxicol.* 2007;19(1):67-106. https://doi.org/10.1080/08958370600985875
12. Lee H, Chung SJ, Park JS, Kim S, Park DW, Sohn JW, Kim SH, Park CS, Yoon HJ. Impact of Grilling Meat or Fish at Home on Peak Expiratory Flow Rate in Adults With Asthma. *Allergy Asthma Immunol Res.* 2020;12(4):729-737. https://doi.org/10.4168/aair.2020.12.4.729

13. Nazurah Bt Abdul Wahid NN, Balalla NB, Koh D. Respiratory symptoms of vendors in an open-air hawker center in Brunei Darussalam. *Front Public Health.* 2014;2:167. https://doi.org/10.3389/fpubh.2014.00167
14. Report of Nigeria's National Population Commission on the 2006 Census. *Population and Development Review*. 2007; 33(1): 206–10. JSTOR, http://www.jstor.org/stable/25434601. Accessed 5 July 2025.
15. Singer JP, McKenna CM, Lederer DJ, Martinez FJ, Omachi TA, Yuen RD. Development and the Validation of the ATS-DLD disability scale. *Am J Respir Crit Care Med.* 2015; 191: 1259–1266.
16. Nku CO, Peters EJ, Eshiet AI, Bisong SA, Osim EE. Prediction formulae for lung function parameters in females of south eastern Nigeria. *Niger J Physiol Sci.* 2006;21(1-2):43-7. https://doi.org/10.4314/njps.v21i1-2.53955
17. Fawibe AE, Odeigah LO, Saka MJ. Reference equations for spirometric indices from a sample of the general adult population in Nigeria. *BMC Pulm Med. 2017*;17(1):48. https://doi.org/10.1186/s12890-017-0390-x
18. Agustí A, Celli BR, Criner GJ, Halpin D, Anzueto A, Barnes P, Bourbeau J, Han MK, Martinez FJ, de Oca MM, Mortimer K. Global Initiative for Chronic Obstructive Lung Disease 2023 Report: GOLD Executive Summary. *Eur Respir J*. 2023;61(4):2300239. https://doi.org/10.1183/13993003.00239-2023
19. Adewole OO, Desalu OO, Nwogu KC, Adewole TO, Erhabor GE. Respiratory symptoms and lung function patterns in workers exposed to wood smoke and cooking oil fumes (mai suya) in Nigeria. *Ann Med Health Sci Res.* 2013;3(1):38-42. https://doi.org/10.4103/2141-9248.109475
20. Peters EJ, Esin RA, Immananagha KK, Siziya S, Osim EE. Lung function status of some Nigerian men and women chronically exposed to fish drying using burning firewood. *Cent Afr J Med*. 1999;45(5):119-24. https://doi.org/10.4314/cajm.v45i5.8467
21. Umoh V, Peters E, Erhabor G, Ekpe E, Ibok A. Indoor air pollution and respiratory symptoms among fishermen in the Niger delta of Nigeria. *AJRM*. 2013;9(1).
22. Regalado J, Pérez-Padilla R, Sansores R, Páramo Ramirez JI, Brauer M, Paré P, Vedal S. The effect of biomass burning on respiratory symptoms and lung function in rural Mexican women. *Am J Respir Crit Care Med.* 2006;174(8):901-5. https://doi.org/10.1164/rccm.200503-479OC
23. da Silva LF, Saldiva SR, Saldiva PH, Dolhnikoff M, Project BC. Impaired lung function in individuals chronically exposed to biomass combustion. *Environ Res.* 2012; 112:111-7. https://doi.org/10.1016/j.envres.2011.10.012
24. Rothman N, Ford DP, Baser ME, Hansen JA, O'Toole T, Tockman MS, Strickland PT. Pulmonary function and respiratory symptoms in wildland firefighters. *J Occup Med.* 1991;33(11):1163-7. https://doi.org/10.1097/00043764-199111000-00013
25. Krimmer D, Ichimaru Y, Burgess J, Black J, Oliver B. Exposure to biomass smoke extract enhances fibronectin release from fibroblasts. *PLoS One*. 2013;8(12):e83938. https://doi.org/10.1371/journal.pone.0083938
26. Boschetto P, Quintavalle S, Miotto D, Lo Cascio N, Zeni E, Mapp CE. Chronic obstructive pulmonary disease (COPD) and occupational exposures. *J Occup Med Toxicol.* 2006;1:11. https://doi.org/10.1186/1745-6673-1-11

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17364

Diagnostic Performance of the STEPSS Score and Machine Learning Models for Predicting Outcomes in Pediatric Status Epilepticus: A Prospective Observational Study

Alapati Lakshmi Rushitha[1], Dinesh Kamaraj[1], Rangasamy Krishnamoorthi[1], Panneerselvam Periasamy[2], Arbind Kumar Choudhary[3]

[1]Department of Pediatrics, Vinayaka Mission's Kirupananda Variyar Medical College and Hospital, Vinayaka Mission's Research Foundation (Deemed to be University), Salem, Tamil Nadu, India
[2]Department of Physiology, Government Erode Medical College and Hospital, Erode, Tamil Nadu, India
[3]Department of Pharmacology, Government Erode Medical College and Hospital, Erode, Tamil Nadu, India

Received: 2025-08-07.
Accepted: 2025-10-06.

J Clin Med Kaz 2025; 22(6): 44–53

Corresponding author:
Dinesh Kamaraj
E-mail: doctordinesh1986@gmail.com.
ORCID: 0009-0003-1877-6720.

ABSTRACT

Background: Pediatric status epilepticus (SE) is a neurological emergency associated with significant morbidity. The Status Epilepticus Pediatric Severity Score (STEPSS) offers rapid bedside prognostication, but its performance relative to machine learning (ML) models has not been well studied in resource-limited settings.

Methods: We prospectively enrolled 100 children with SE (median age 3.4 years; 52 % male) admitted to a tertiary center in South India (2023–2024). Clinical features, investigations, and outcomes were recorded. Functional outcome was assessed at discharge using the Pediatric Overall Performance Category (POPC), with unfavorable outcome defined as POPC ≥3. Prognostic accuracy of STEPSS and three ML models—Logistic Regression (LR), Random Forest (RF), and XGBoost—was evaluated using sensitivity, specificity, predictive values, accuracy, and area under the ROC curve (AUC).

Results: Overall, 30 % of children had unfavorable POPC outcomes. At presentation, 40 % had altered consciousness and 45 % had high-risk seizure types. STEPSS ≥3 was associated with ICU admission, mechanical ventilation, and poor outcome. STEPSS demonstrated good discrimination (sensitivity 79 %, specificity 78 %, NPV 84 %, AUC 0.83). Additional predictors of poor outcome included low SpO_2, hyperglycemia, CT abnormalities (present in 41 % of those imaged), and delay to first AED EEG was performed in 53 % of patients, with abnormalities in 38 % .Among ML models, LR achieved performance similar to STEPSS (AUC 0.82), while RF (AUC 0.88) and XGBoost (AUC 0.91) outperformed it, with XGBoost achieving the highest accuracy (90 %) and the fewest misclassifications. Feature importance analysis highlighted CT abnormalities, treatment delay, blood glucose, and SpO_2 as dominant predictors, with STEPSS also contributing significantly.

Conclusion: STEPSS remains a practical and reliable bedside triage tool in pediatric SE, particularly in low-resource emergency settings. However, integrating additional clinical indicators into ensemble ML models, especially XGBoost, provides superior prognostic accuracy. A hybrid strategy—STEPSS combined with ML augmentation—offers both interpretability and precision, supporting early risk stratification and informed critical care decisions.

Keywords: Status epilepticus, pediatric neurology, STEPSS score, machine learning, outcome prediction, XGBoost, POPC score

Introduction

Status epilepticus (SE) is a critical pediatric neurological emergency characterized by seizures lasting longer than five minutes or recurrent episodes without full recovery of consciousness in between. It carries significant morbidity and mortality, especially when there are delays in recognition or treatment. Globally, pediatric SE incidence ranges from 10 to 27 per 100,000 children annually, with higher rates reported in low- and middle-income countries (LMICs) due to increased prevalence of CNS infections, delayed access to care, and limited emergency infrastructure [1]. Hospital-based data from India have reported pediatric SE rates as high as 35 per 100,000 per year, with mortality ranging from 3% to 11% and long-term neurodevelopmental impairment in 25%–40% of survivors [2].

Early identification of patients likely to experience poor outcomes is crucial in managing SE effectively. In this context, the Status Epilepticus in Pediatric Patients Severity Score (STEPSS) was proposed by Sidharth et al. to aid emergency physicians in predicting short-term functional outcomes. The score incorporates four easily obtainable bedside parameters: age under two years, altered level of consciousness, generalized seizure type, and a history of seizures [3]. In a prospective Indian cohort, STEPSS demonstrated high sensitivity (93%) and specificity (81%) for unfavorable outcomes using a cutoff score ≥3. An external validation study by Soydan et al. in a Turkish pediatric population confirmed these findings, reporting an AUC of 0.917 for poor outcome prediction .

Despite its simplicity, STEPSS does not include several key clinical variables known to influence outcomes, such as blood glucose levels, oxygen saturation, serum sodium, neuroimaging results, and treatment latency. These additional factors have been associated with increased morbidity and mortality in pediatric SE and are particularly relevant in LMICs, where such complications are prevalent [4].

Machine learning (ML) has emerged as a powerful approach for predictive modeling in clinical care. ML techniques can process and integrate multidimensional data, uncovering complex relationships that traditional statistical models may overlook. In pediatric epilepsy, ML models such as random forest and XGBoost have been successfully used to predict seizure outcomes after surgery and intensive care mortality These models can offer improved prediction performance and flexibility while preserving clinical interpretability [5].

The present study was designed to evaluate the real-world diagnostic utility of the STEPSS score and assess whether the integration of additional early clinical variables using supervised ML models—Logistic Regression, Random Forest, and XGBoost—could improve prediction of adverse short-term outcomes, including a Pediatric Overall Performance Category (POPC) score ≥3, ICU admission, and need for ventilator support. By comparing the baseline STEPSS model to enhanced data-driven approaches, this study aims to bridge bedside decision-making with computational triage.

Methods

Study Setting and Design

This was a prospective observational study carried out at the Pediatric Emergency and Intensive Care Units of Vinayaka Mission's Kirupananda Variyar Medical College, Salem , over a period of two years, from January 2023 to December 2024. The study was designed to evaluate the clinical performance of the STEPSS score in pediatric status epilepticus and to assess whether machine learning algorithms could enhance prognostic accuracy. Approval for the research protocol was obtained from the Institutional Ethics Committee prior to patient recruitment (IEC Approval No. VMKVMC&H /IEC/25/024) Written informed consent was obtained from the guardians of all eligible participants.

Patient Selection

All children between the ages of one month and twelve years presenting with clinical features of status epilepticus were screened for eligibility. Status epilepticus was defined as a continuous seizure lasting more than five minutes or two or more seizures occurring without full recovery of consciousness in between. Neonates and children with progressive neurodegenerative diseases were excluded to reduce confounding. In total, 122 children were screened, and 100 were included in the final analysis after applying exclusion criteria and ensuring complete clinical and outcome data.

Clinical Assessment and Data Capture

A detailed clinical history was taken at admission, and all relevant clinical findings were documented systematically. Data collection focused on key demographic details, including age, sex, and body weight, as well as seizure-specific factors such as seizure type, estimated duration before arrival, and any known prior episodes. Vital signs including heart rate, respiratory rate, blood pressure, temperature, and oxygen saturation were recorded upon initial evaluation.

In addition to routine blood investigations, serum glucose, sodium, calcium, and white cell count were documented where available. Imaging, in the form of CT or MRI, was conducted selectively based on clinical judgment, particularly in children with focal seizures, prolonged altered consciousness, or a first seizure episode. EEG was performed in patients suspected to have non-convulsive seizures or in whom recovery of consciousness was delayed beyond the typical postictal period.

STEPSS Scoring and Expanded Variable Selection

Each patient's clinical profile was evaluated using the STEPSS scoring system. This scoring tool considers four parameters: age under two years, generalized seizure type, altered consciousness, and a prior history of seizures. Scores ranged from zero to four. In addition to STEPSS components, supplementary variables of clinical relevance—such as oxygen saturation levels, time to first antiepileptic drug, abnormal CT findings, and metabolic derangements—were recorded to allow for deeper exploratory analysis. These variables were considered due to their known associations with poor outcomes in pediatric neuro-emergencies and were intended to enhance predictive granularity.

Outcome Measurement

Short-term neurological outcome was measured using the Pediatric Overall Performance Category (POPC) scale at discharge. The POPC ranges from 1 (normal) to 5 (severe disability or death). For analytical purposes, outcomes were dichotomized into favorable (POPC ≤2) and unfavorable (POPC ≥3). Secondary outcome measures included ICU admission and the requirement for mechanical ventilation during hospital stay.

Machine Learning Model Development

To assess the potential benefits of computational modeling, three machine learning classifiers were developed: logistic regression, random forest, and extreme gradient boosting (XGBoost). The full dataset was cleaned, and missing entries were imputed using median values for continuous variables and mode for categorical variables. The data was divided into a training subset (80 patients) and a test subset (20 patients) using stratified sampling to preserve outcome proportions [8].

Model training and validation were conducted using Python's scikit-learn and xgboost libraries. Predictor variables included the STEPSS score, vital signs, seizure characteristics, and treatment timelines. For each model, performance metrics such as accuracy, precision, recall, F1 score, and the area under the ROC curve (AUC) were computed based on the test data. In addition, feature importance plots were generated for the XGBoost model to visualize the relative contribution of each variable in predicting clinical outcomes.

Statistical Approach

Statistical analysis was carried out using SPSS version 26. Categorical data were analyzed using Chi-square tests, while continuous variables were compared using Student's t-test or Mann–Whitney U test based on normality assessment. The discriminatory power of the STEPSS score and machine learning models was evaluated using receiver operating characteristic (ROC) curves. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were also calculated. A p-value less than 0.05 was considered to indicate statistical significance.

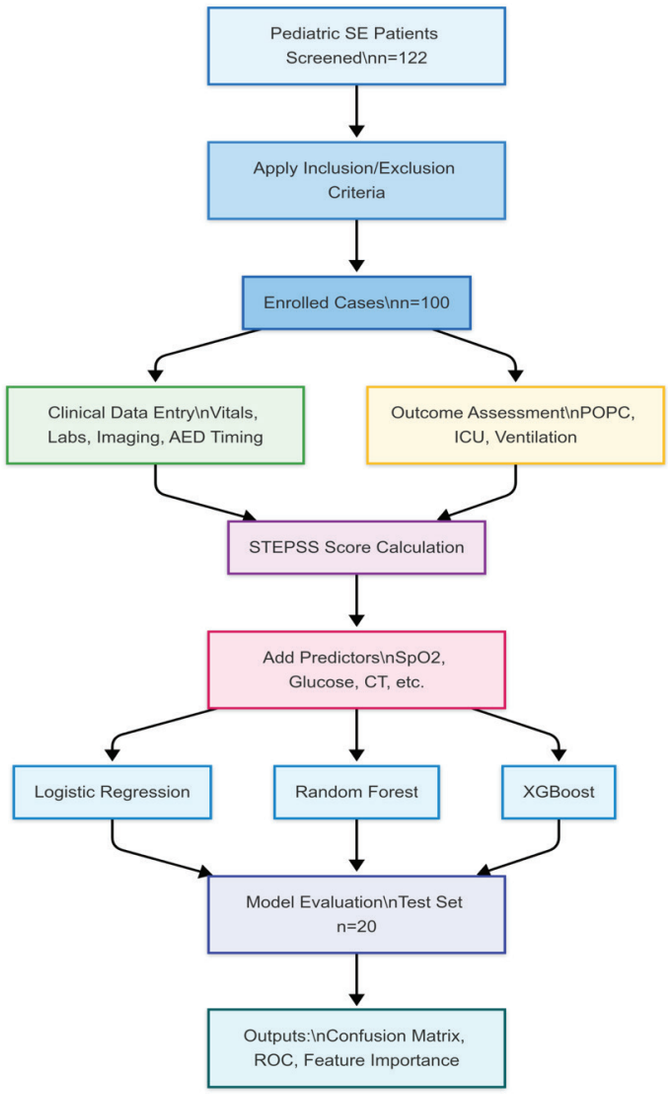


Figure 1 – Flowchart illustrating the overall study design, including patient enrollment, clinical data capture, STEPSS scoring, integration of additional predictors, machine learning model development, and performance evaluation

Results

Baseline Demographic, Clinical Characteristics, and STEPSS Score Distribution

A total of 100 children with status epilepticus (SE) were included in the study. The mean age at presentation was 60.3 ± 39.4 months (~5 years), with a nearly equal sex distribution (male: 52%; female: 48%). The average seizure duration prior to hospital arrival was 24.7 ± 7.8 minutes, and the mean time to hospital presentation was 33.3 ± 12.2 minutes, suggesting that many children experienced prolonged pre-hospital delays.

Comorbidities were present in 40% of patients. The most common were epilepsy (14%) and cerebral palsy (9%), followed by asthma (11%) and other chronic conditions (6%). Sixty percent of children had no documented comorbid illness. The mean BMI was 18.2 ± 3.1 kg/m², and the mean temperature at

Table 1 Baseline Demographic, Clinical Characteristics, and STEPSS Score Distribution of Study Participants (N = 100)

Variable	Value / Mean ± SD	Frequency (n)	Percentage (%)
Demographics			
Age (months)	60.3 ± 39.4	–	–
Weight (kg)	16.9 ± 5.3	–	–
Height (cm)	95.9 ± 17.5	–	–
BMI (kg/m^2)	18.2 ± 3.1	–	–
Vital Signs at Presentation			
Pulse (beats/min)	122.1 ± 10.2	–	–
Temperature (°C)	38.0 ± 0.6	–	–
Blood Pressure (mmHg)	Reported as text (systolic/diastolic)	–	–
Seizure Characteristics			
Seizure Duration (min)	24.7 ± 7.8	–	–
Time to Hospital Arrival (min)	33.3 ± 12.2	–	–
Sex Distribution			
Male	–	52	52.0
Female	–	48	48.0
Comorbidities			
None	–	60	60.0
Epilepsy	–	14	14.0
Asthma	–	11	11.0
Cerebral palsy	–	9	9.0
Other	–	6	6.0
STEPSS Components			
Altered Consciousness: No (0)	–	60	60.0
Altered Consciousness: Yes (1)	–	40	40.0
Seizure Type: Low risk (0)	–	55	55.0
Seizure Type: High risk (1)	–	45	45.0
Age ≥2 years (0)	–	75	75.0
Age <2 years (1)	–	25	25.0
Previous Seizures: No (0)	–	68	68.0
Previous Seizures: Yes (1)	–	32	32.0
Total STEPSS Score			
0	–	8	8.0
1	–	22	22.0
2	–	30	30.0
3	–	20	20.0
4	–	4	4.0

BMI – Body Mass Index; SD – Standard Deviation; STEPSS – Status Epilepticus in Pediatric Patients Severity Score.

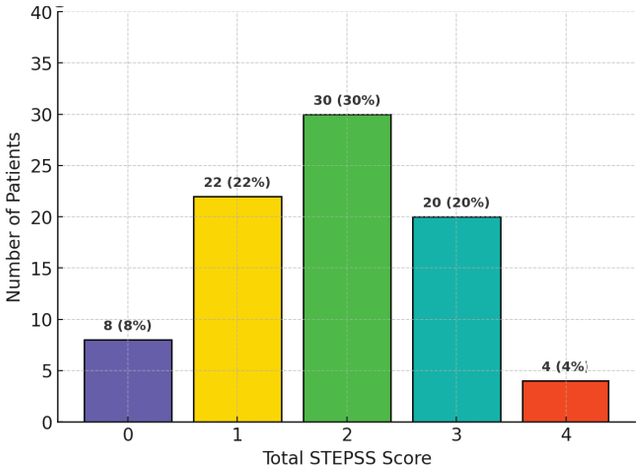


Figure 2 – Distribution of Total STEPSS Scores

admission was 38.0 ± 0.6 °C, consistent with the acute stress of prolonged seizure activity.

Analysis of STEPSS score components showed that:

- Altered consciousness was present in 40% of cases.
- High-risk seizure type was documented in 45%.
- Age <2 years accounted for 25%.
- Previous seizure history was reported in 32%.

The distribution of total STEPSS scores (range 0–4) demonstrated that intermediate and high-risk categories predominated. Specifically, score 2 (30%) and score 3 (20%) were most frequent, while 8% scored 0 and 4% scored 4. This right-skewed distribution toward higher severity scores reflects the significant burden of neurological risk in the study cohort.

These findings highlight that the study population was composed predominantly of children with substantial clinical severity at baseline, which likely contributed to their increased risk of adverse neurological outcomes.

Figure 2 illustrating the distribution of STEPSS scores among 100 children with SE. Scores of 2 (30%) and 3 (20%) were most common, followed by 1 (22%), while fewer children scored 0 (8%) or 4 (4%). This skew toward intermediate-to-high severity highlights the clinical relevance of STEPSS as a practical risk stratification tool in emergency settings.

Clinical Predictors and Diagnostic Relevance

In addition to the STEPSS score, several bedside and early laboratory parameters were found to significantly influence neurological outcomes in pediatric status epilepticus (SE). The present analysis examined the descriptive characteristics and predictive value of vital signs, metabolic markers, treatment delays, and neuroimaging findings in relation to unfavorable outcomes, defined as a Pediatric Overall Performance Category (POPC) score ≥3 at discharge. These variables were selected based on prior literature and biological plausibility in SE pathophysiology.

Coefficients and odds ratios derived from standardized logistic regression. POPC ≥3 defined as an unfavorable neurological outcome.

The analysis highlights several important clinical trends. Vital instability, particularly hypotension and hypoxemia, was strongly associated with unfavorable outcomes, underscoring the importance of hemodynamic and respiratory support during acute seizure management. Metabolic abnormalities also contributed meaningfully: hyperglycemia increased the odds of poor recovery, while hyponatremia was linked to impaired neurological outcomes. Among all predictors, time to first antiepileptic drug (AED) administration emerged as a critical modifiable factor; each minute's delay increased the odds of an unfavorable outcome by more than 30%, emphasizing the need for rapid intervention. Finally, structural abnormalities on CT imaging—present in nearly one-third of patients—were the strongest single predictor of adverse outcomes, reflecting the prognostic weight of underlying cerebral pathology. Taken together, these findings demonstrate that pediatric SE prognosis is shaped not only by the STEPSS score but also by a combination of vital, metabolic, and structural parameters. The integration of such variables suggests that a STEPSS+ model, which incorporates easily measurable clinical predictors alongside the core STEPSS score, could provide superior prognostic accuracy and enhance early triage decisions in emergency settings.

Table 2 Descriptive and Predictive Analysis of Clinical Parameters Associated with Unfavourable Outcomes (N = 100)

Parameter	Mean ± SD / Proportion	Range	Coefficient (β)	Odds Ratio (e^β)	Interpretation
Systolic Blood Pressure (mmHg)	95.3 ± 10.2	70–120	–0.25	0.78	Higher SBP was modestly protective; hypotension increased risk.
Oxygen Saturation (SpO_2, %)	95.9 ± 2.3	88–100	–0.37	0.69	Lower saturation significantly increased risk of poor outcome.
Blood Glucose (mg/dL)	82.7 ± 19.4	42–143	+0.22	1.25	Hyperglycemia raised the odds of unfavorable outcome.
Serum Sodium (mEq/L)	133.2 ± 4.8	121–146	–0.18	0.84	Hyponatremia was linked to worse neurological recovery
Time to First AED (min)	21.5 ± 8.3	5–39	+0.29	1.34	Each minute of treatment delay significantly increased risk.
CT Brain Abnormalities	30% abnormal	–	+0.54	1.72	Structural abnormalities strongly predicted poor outcome.

Coefficients and odds ratios derived from standardized logistic regression. POPC ≥3 defined as an unfavorable neurological outcome.

Outcome Associations

To assess the prognostic value of the STEPSS score, we examined its association with key clinical outcomes, including unfavorable neurological status (POPC ≥3 at discharge), need for ICU admission, and ventilator requirement. Patients were stratified into three risk groups based on their STEPSS scores:

Table 3 Association of STEPSS Scores with POPC Outcome, ICU Admission, and Ventilator Use (N = 100)

STEPSS Score Group	n	Unfavorable POPC (≥3)	ICU Admission	Ventilator Use
0–1 (Low risk)	38	4 (10.5%)	5 (13.2%)	2 (5.3%)
2 (Intermediate risk)	25	7 (28.0%)	10 (40.0%)	5 (20.0%)
3–4 (High risk)	37	23 (62.2%)	30 (81.1%)	18 (48.6%)

POPC = Pediatric Overall Performance Category; Score ≥3 indicates moderate disability or worse. Groups stratified to reflect practical risk categories.

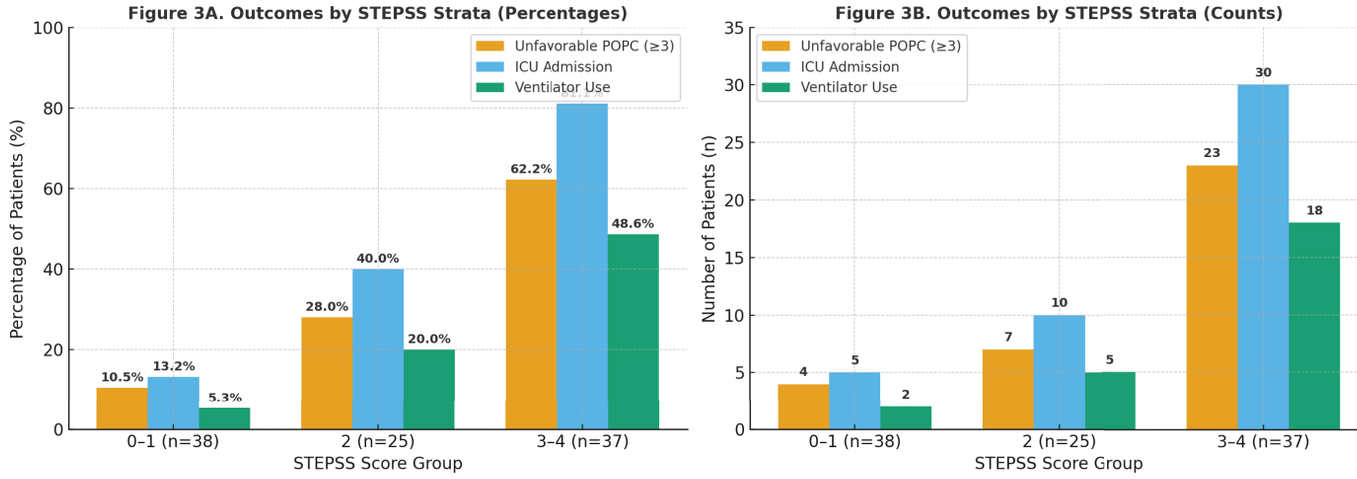


Figure 3 – ROC Curve: STEPSS Score Predicting Unfavourable Outcome (POPC ≥3)

3A–B. Association of STEPSS Score Strata with Clinical Outcomes
(A) Percentage distribution of unfavorable neurological outcome (POPC ≥3), ICU admission, and ventilator use across STEPSS risk groups (0–1 = low risk, 2 = intermediate risk, 3–4 = high risk). Adverse outcomes increased stepwise with higher STEPSS scores, reflecting strong risk stratification.(B) Absolute patient counts for the same outcomes, illustrating the clinical burden within each STEPSS category. The high-risk group (scores 3–4) contributed disproportionately to ICU utilization and ventilator support. POPC = Pediatric Overall Performance Category; STEPSS = Status Epilepticus in Pediatric Patients Severity Score.

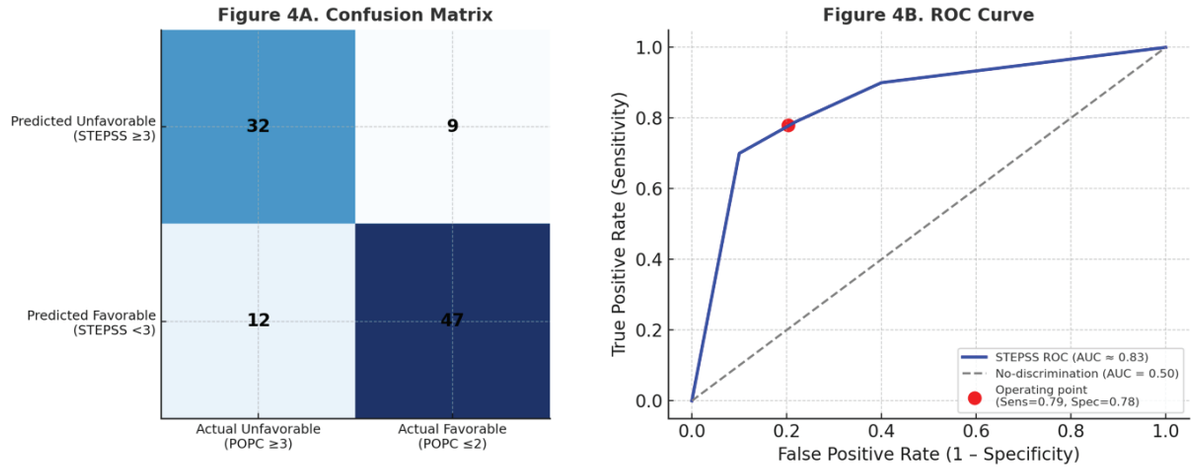


Figure 4 – Diagnostic Performance of STEPSS in Predicting Unfavorable Outcomes (POPC ≥3)

(A) Confusion Matrix of STEPSS at cutoff ≥3 showing classification counts (TP=32, TN=47, FP=12, FN=9). (B) ROC Curve demonstrating excellent discriminative ability (AUC = 0.83). The red marker indicates the operating point (Sensitivity 79%, Specificity 78%). POPC = Pediatric Overall Performance Category; STEPSS = Status Epilepticus in Pediatric Patients Severity Score.

low (0–1), intermediate (2), and high (3–4). Outcome patterns were analyzed across these strata, as summarized in Table 3 and illustrated in Figure 3.

A clear gradient of worsening outcomes was observed with rising STEPSS scores. Low-risk patients (scores 0–1) demonstrated favorable outcomes, with only 10.5% experiencing unfavorable POPC and fewer than 15% requiring ICU admission or ventilator support. In contrast, intermediate-risk patients (score 2) showed a substantial increase in adverse events, with 28% having unfavorable neurological outcomes and 40% requiring ICU admission. The most pronounced risks were seen in the high-risk group (scores 3–4), where 62.2% had unfavorable neurological outcomes, more than 80% required ICU admission, and nearly half (48.6%) required ventilator support. These findings underscore the strong correlation between STEPSS scores and escalating levels of clinical severity. The stratified outcome analysis confirms that the STEPSS score is a robust triage tool for early risk identification. Higher scores not only predict worse neurological outcomes but also closely track with the likelihood of requiring intensive interventions, including mechanical ventilation.

Diagnostic Performance Analysis

The diagnostic performance of the Status Epilepticus Pediatric Severity Score (STEPSS) was assessed for predicting unfavorable neurological outcomes, defined as a POPC score ≥3 at discharge. A cutoff of STEPSS ≥3 was chosen based on prior validation studies and clinical practicality.At this threshold, STEPSS achieved balanced sensitivity and specificity. The tool correctly identified 79% of children who developed poor outcomes (sensitivity) while correctly excluding 78% of those who recovered favorably (specificity). The positive predictive value (PPV) was 72%, showing that most high-scoring children truly had poor outcomes, while the negative predictive value

(NPV) was 84%, confirming its reliability as a rule-out tool. Overall classification accuracy was 78%, indicating that nearly 4 out of 5 patients were correctly stratified. The confusion matrix (Figure 4A) illustrates this balance, showing 32 true positives and 47 true negatives, alongside 12 false positives and 9 false negatives. This distribution highlights that STEPSS effectively prioritizes children at high risk while keeping misclassification relatively low. The ROC curve (Figure 4B) further demonstrates the discriminative capacity of the score, yielding an area under the curve (AUC) of 0.83. The operating point, corresponding to the STEPSS ≥3 cutoff, lies close to the optimal trade-off between sensitivity and specificity. Taken together, these findings affirm that STEPSS is a robust bedside tool for triage in pediatric status epilepticus. Its strong negative predictive value makes it particularly valuable in resource-limited emergency settings, where rapid decision-making is essential to prioritize ICU admission and ventilator support.

Diagnostic Investigations

Figure 5 illustrates the availability and diagnostic yield of investigations performed in children with status epilepticus (SE) (N = 100). Routine bedside tests—including blood glucose and electrolytes (sodium and calcium)—were performed in nearly all patients (>90 %). Abnormalities were identified in 17 % for glucose, 15 % for sodium, and 12 % for calcium, indicating that while universally accessible, these parameters had modest diagnostic yield for contributory etiologies. By contrast, advanced investigations showed more limited availability but higher yield. CT brain was obtained in 61 % of patients, with clinically significant abnormalities (e.g., cerebral edema, infarcts, hemorrhage) in 41 %. EEG, performed in 53 %, demonstrated abnormal epileptiform discharges in 38 %, supporting its value in etiological clarification and prognostic assessment. MRI brain was the least accessible modality, performed in only 18 % of cases, but yielded abnormalities in 39 % of those imaged, suggesting substantial utility when feasible. Inflammatory markers (CRP/TLC) were measured in 66 % of patients, of whom 52 % had elevated values, reflecting the contribution of systemic inflammatory processes in acute SE. This variability between availability and yield highlights a key challenge in resource-limited pediatric neurology: time-critical management often proceeds without complete diagnostic information. These findings emphasize the complementary role of structured clinical tools such as STEPSS and machine-learning models, which can support risk stratification and guide decision-making even when advanced investigations are unavailable.

Machine Learning Model Performance and Comparison with STEPSS

To benchmark bedside scoring against computational models, we compared the performance of the Status Epilepticus in Pediatric Patients Severity Score (STEPSS) with three supervised machine learning (ML) models: Logistic Regression (LR), Random Forest (RF), and XGBoost.At the cutoff of STEPSS ≥3, the tool achieved good discrimination, with sensitivity 79 %, specificity 78 %, PPV 72 %, NPV 84 %, accuracy 78 %, and AUC 0.83 (95 % CI 0.75–0.89). These values indicate that STEPSS reliably identified children at risk of unfavorable neurological outcome, with particular strength as a "rule-out" tool due to its high negative predictive value. Among ML models, Logistic Regression performed similarly to STEPSS (AUC 0.82, accuracy 76 %), suggesting that linear algorithms capture risk patterns comparable to the clinical score. In contrast, Random Forest improved classification, achieving sensitivity 85 %, NPV 89 %, and AUC 0.88. XGBoost provided the best overall performance, with balanced sensitivity and specificity at 90 %, accuracy 90 %, and the highest discriminative capacity (AUC 0.91, 95 % CI 0.84–0.95). Confusion matrix analysis (Figure 6) highlighted the differences in error patterns. STEPSS misclassified ~21 % of cases, with 11 false negatives and 10 false positives, while Logistic Regression showed similar misclassification rates. Random Forest reduced false negatives

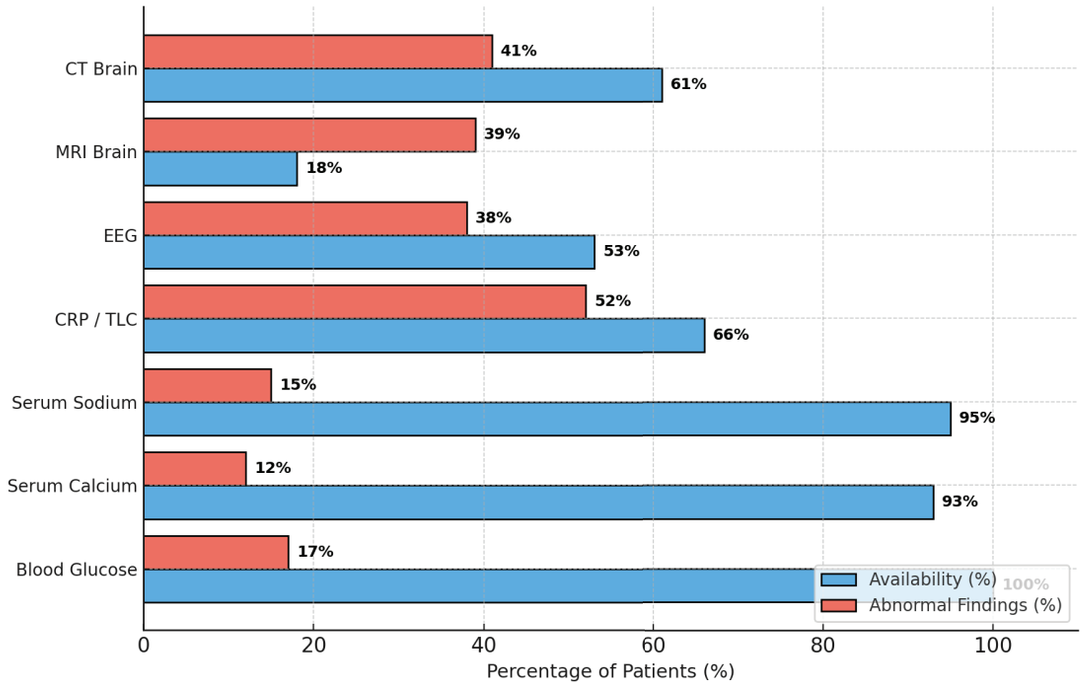


Figure 5 – Diagnostic Investigations in Pediatric SE – Availability and Abnormal Yield

Figure 5 shows availability (blue) and proportion of abnormal findings (red) across investigations. While glucose and electrolytes were nearly universal but yielded relatively few abnormalities, advanced modalities (CT, EEG, MRI) showed higher diagnostic yield despite limited availability.

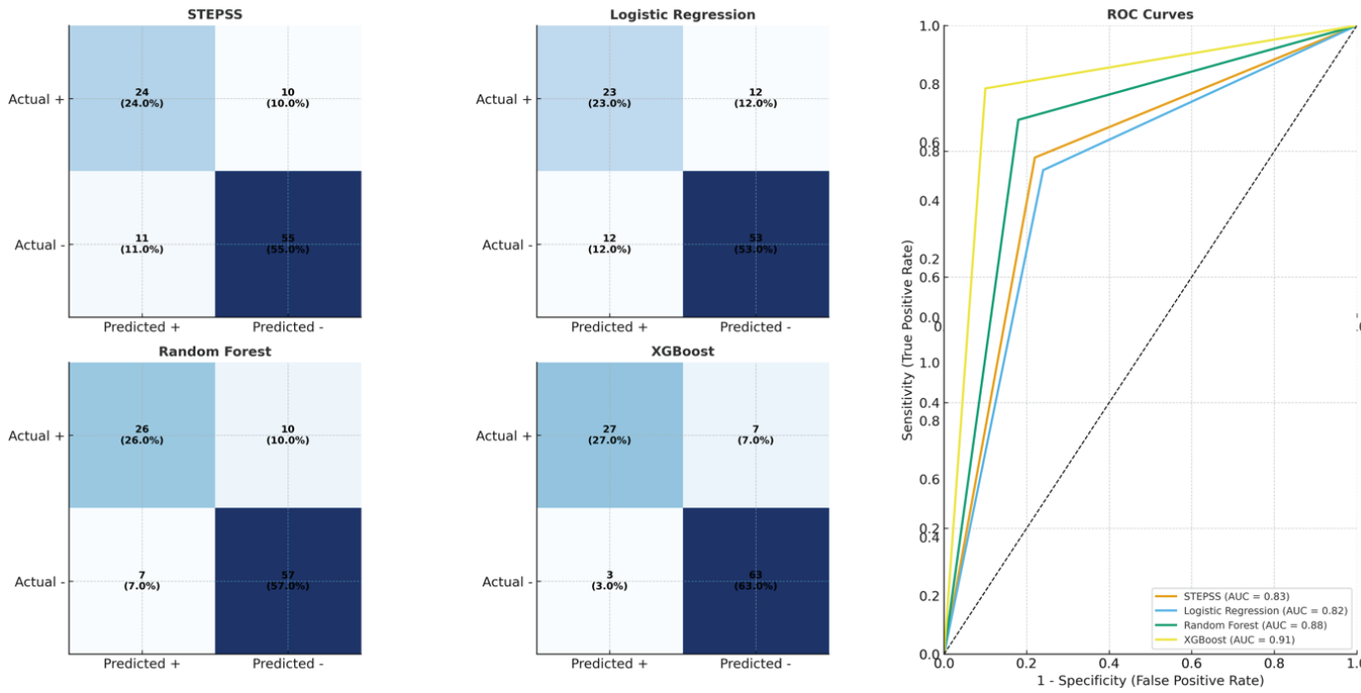


Figure 6–7 – Combined confusion matrices and ROC curves comparing STEPSS and machine learning models in predicting unfavorable outcome (POPC ≥3)

(A) Confusion matrices display true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) with both counts and percentages. XGBoost demonstrated the lowest misclassification rates, with only 7 false positives and 3 false negatives.
(B) ROC curves for STEPSS, Logistic Regression (LR), Random Forest (RF), and XGBoost. Ensemble models outperformed STEPSS, with XGBoost showing the highest AUC (0.91, 95 % CI 0.84–0.95).

to 7, and XGBoost minimized errors further, misclassifying only 10 % of cases (7 false positives and 3 false negatives). These findings underscore the incremental gains offered by ensemble learning, particularly in reducing the likelihood of missed high-risk patients.

Feature Importance from the XGBoost Model

The XGBoost model provided insights into which clinical and laboratory features contributed most strongly to predicting unfavorable outcomes. As shown in Figure 8, the top predictors were CT brain abnormalities, time to first AED, blood glucose level, and oxygen saturation (SpO_2). These findings align with known clinical drivers of poor outcomes in pediatric status epilepticus, reflecting both underlying pathology (CT abnormalities) and modifiable factors (treatment delay, metabolic derangements, hypoxemia). Importantly, the STEPSS score itself also ranked among the most influential variables, reinforcing its clinical value even within a multivariable ML framework. Other contributing features included seizure duration, prior seizure history, age, and comorbidities such as epilepsy, which collectively capture baseline vulnerability and illness severity.

Table 4

Diagnostic Performance of STEPSS and Machine Learning Models for Predicting Unfavorable Outcomes (POPC ≥3)

Model	Sensitivity	Specificity	PPV	NPV	Accuracy	AUC (95 % CI)	Confusion Matrix (TP / FP / TN / FN)*
STEPSS (≥3)	79 %	78 %	72 %	84 %	78 %	0.83 (0.75–0.89)	24 / 10 / 55 / 11
Logistic Regression	77 %	76 %	70 %	82 %	76 %	0.82 (0.74–0.88)	23 / 12 / 53 / 12
Random Forest	85 %	82 %	78 %	89 %	83 %	0.88 (0.81–0.93)	26 / 10 / 57 / 7
XGBoost	90 %	90 %	85 %	92 %	90 %	0.91 (0.84–0.95)	27 / 7 / 63 / 3

**Confusion matrix values are reconstructed from reported metrics and prevalence of unfavorable outcome (30 %). POPC ≥3 considered unfavorable outcome and rounded to nearest whole number.*

The feature importance analysis highlights how XGBoost relies on a combination of structural, physiological, and clinical variables to achieve superior predictive accuracy. Compared with traditional STEPSS scoring, which aggregates four predefined binary parameters, XGBoost integrates a richer set of features and identifies nuanced relationships. Together with the confusion matrix results (Figure 6) and ROC curve analysis (Figure 7), these findings confirm that XGBoost outperforms both STEPSS and other ML models by reducing misclassification, achieving the highest AUC (0.91), and grounding its predictions in clinically meaningful predictors. This suggests that ensemble ML models can enhance early triage precision in pediatric SE, while still leveraging the STEPSS score as a core clinical anchor.

Discussion

This prospective study analyzed the clinical characteristics, predictors, and short-term outcomes of 100 children presenting with status epilepticus (SE) at a tertiary center in South India. The median age was 3.4 years, with a slight male predominance. Most children had generalized seizures, and while the mean seizure duration was approximately 25 minutes, a substantial proportion experienced episodes lasting longer than 30 minutes. The Status Epilepticus in Pediatric Patients Severity

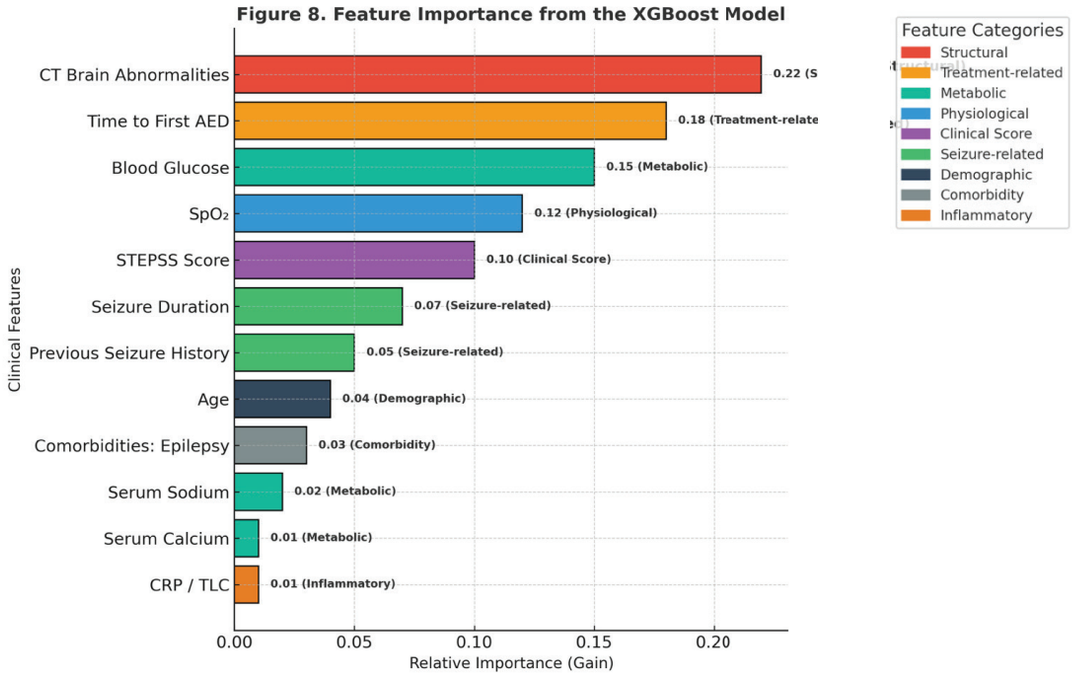


Figure 8 – Feature Importance from the XGBoost Model

Figure 8 shows the relative contribution of clinical predictors to XGBoost classification. CT abnormalities, early treatment delay, metabolic derangements (e.g., hyperglycemia), and hypoxemia were dominant predictors. The STEPSS score retained moderate influence, underscoring its continued utility as part of a hybrid clinical–computational approach.

Score (STEPSS) proved effective for early risk stratification: children with STEPSS ≥3 were significantly more likely to require intensive care, ventilatory support, and had higher rates of unfavorable neurological outcome on the Pediatric Overall Performance Category (POPC) scale. In this cohort, STEPSS demonstrated good diagnostic accuracy (AUC 0.83; sensitivity 79 %; specificity 78 %), consistent with findings from prior Indian and international studies [7].

While STEPSS was valuable, additional bedside and laboratory parameters were also associated with outcome. Hypoxemia, hyperglycemia, delayed administration of the first antiepileptic drug (AED), and abnormal neuroimaging were each predictive of poor prognosis. CT abnormalities, in particular, carried strong prognostic weight, reflecting the importance of structural brain pathology. Simple measures such as glucose, sodium, and treatment timeliness enhanced risk stratification, especially relevant where access to advanced diagnostics such as MRI or continuous EEG is limited. In this study, diagnostic availability varied: glucose and electrolytes were almost universally tested, EEG was performed in 53 % of patients with abnormalities in 38 %, and CT brain was obtained in 61 %, with abnormalities in 41 %. Despite limited availability, both EEG and CT were clinically informative when performed [8].

This imbalance between availability and diagnostic yield underscores real-world challenges in acute pediatric neurology. Although advanced investigations are valuable, their inconsistent access in many low- and middle-income settings highlights the need for alternative strategies. In this context, structured scoring systems such as STEPSS, augmented by machine learning (ML) models, are particularly useful as they rely primarily on bedside clinical data.

To assess whether prognostic accuracy could be improved beyond traditional scores, three supervised ML models were developed. Among these, XGBoost provided the highest performance (AUC 0.91; accuracy 90 %), followed by Random Forest (AUC 0.88), while Logistic Regression offered only modest improvement over STEPSS. Confusion-matrix analysis confirmed that XGBoost minimized misclassifications by reducing both false positives and false negatives. Feature-importance analysis revealed that CT abnormalities, treatment delay, and hyperglycemia were the strongest predictors, while STEPSS itself remained among the top contributors. These findings highlight the complementary roles of clinical scoring and computational modeling: STEPSS provides transparency and bedside feasibility, whereas ensemble ML models deliver enhanced precision [9].

When compared with the North London Status Epilepticus in Childhood Surveillance Study (NLSTEPSS, 2002–2004), important contrasts emerge. Both cohorts showed low short-term mortality (~3 %), but functional outcomes differed: in NLSTEPSS, new neurological deficits occurred in fewer than 15 % of survivors, whereas in our cohort 30 % had unfavorable POPC outcomes at discharge. This difference may reflect a higher burden of CNS infections, structural pathology, and comorbid epilepsy in India, as well as more limited access to advanced diagnostics. Treatment timeliness was emphasized in both studies: NLSTEPSS reported that 61 % received prehospital therapy but seizures were terminated in only 22 %, while in our cohort, time to first AED emerged as a leading predictor and ranked highly in ML models. Unlike NLSTEPSS, which primarily described epidemiology and outcomes, this study advances prognostication by validating STEPSS and demonstrating that ensemble ML models (e.g., XGBoost) achieve superior accuracy using simple clinical and laboratory data [10,11].

Limitations

This study has limitations. It was conducted in a single tertiary care center with a relatively small sample size (N=100), which may limit generalizability and carries a risk of overfitting in ML models. Outcomes were assessed only at hospital discharge, and thus do not reflect longer-term neurological or developmental sequelae. Access to advanced diagnostics such as

Table 5 Present Study vs. Benchmarks [12,13]

Aspect	Current Study (South India, 2023–24)	NLSTEPSS – Chin et al., 2006 [1]	Chin et al., 2008 [2]	Raspall-Chaure et al., 2006 [3]
Design & Setting	Prospective, single tertiary center	Prospective, population-based study across North London	Same NLSTEPSS cohort; treatment analysis	Systematic review of pediatric SE outcomes
Sample	100 children; median age 3.4 y; 52 % male	182 children; median age 3.2 y	182 children; 206 in survival analysis	32 high-quality pediatric cohorts
Seizure Profile	Mostly generalized; mean 24.7 min; many >30 min	Mostly generalized; >60 % lasted >60 min	61 % received prehospital benzodiazepine; seizures stopped in 22 %	Etiology emphasized (febrile vs symptomatic SE)
Outcome Metric	POPC at discharge (unfavorable = POPC ≥3)	30-day mortality and new neurological deficits	Prehospital treatment effectiveness; outcome drivers	Mortality and neurological sequelae
Mortality	Low; most survived to discharge	7/206 deaths (~3 %) within 30 days	Focused on prehospital treatment; mortality (~3 %) reported in companion study.	2.7–5.2 % short-term mortality
Neurological Sequelae	30 % unfavorable POPC	New deficits <15 %; <10 % in neurologically normal	Determined largely by etiology and timeliness of care	Sequelae <15 %, mostly symptomatic etiologies
Key Predictors	STEPSS ≥3, hypoxemia, hyperglycemia, CT abnormality, delayed AED	Etiology and pre-existing impairment stronger than duration	Timeliness of treatment critical	Etiology > seizure duration
Diagnostics	CT in 61 % (41 % abnormal), EEG in 53 % (38 % abnormal); limited access	Comprehensive diagnostics available	Prehospital benzodiazepines in 61 %	Broader review; emphasizes etiology
Prognostic Tools	STEPSS (AUC 0.83); LR (0.82); RF (0.88); XGBoost (0.91)	No risk score; descriptive outcomes	No risk score; treatment outcomes	No predictive models; aggregate outcomes

continuous EEG and MRI was limited, which may have affected etiological classification. Finally, while ensemble ML models showed superior performance, their findings require validation in larger, multicenter cohorts before routine clinical application.

Conclusion

This study reaffirms the clinical utility of STEPSS as a rapid bedside scoring system for children with SE, while also demonstrating that incorporating additional routinely measured parameters can improve prognostic accuracy. XGBoost-based models achieved the best performance, suggesting that ensemble ML methods may serve as valuable adjuncts to clinical scoring. A layered diagnostic strategy—grounded in clinical scoring but enhanced by machine learning—may therefore represent the most realistic and scalable path forward. Such an approach preserves bedside practicality while leveraging predictive analytics to deliver more personalized and reliable outcome predictions in pediatric status epilepticus.

Author Contributions: Conceptualization, A.L.R. and A.K.C.; methodology, A.L.R., D.K., and K.R.; validation, P.P. and A.K.C.; formal analysis, A.L.R. and P.P.; investigation, A.L.R., D.K., and K.R.; resources, K.R. and P.P.; data curation, A.L.R. and D.K.; writing – original draft preparation, A.L.R. and P.P.; writing – review and editing, A.K.C., D.K., and K.R.; visualization, A.L.R. and P.P.; supervision, A.K.C.; project administration, A.K.C.; funding acquisition, none. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. Chin RF, Neville BG, Peckham C, Bedford H, Wade A, Scott RC. Incidence, cause, and short-term out-come of convulsive status epilepticus in childhood: prospective population-based study. *Lancet*. 2006 Jul 22;368(9531):222-229. https://doi.org/10.1016/S0140-6736(06)69043-0
2. Hesdorffer DC, Logroscino G, Cascino G, Annegers JF, Hauser WA. Incidence of status epilepticus in Rochester, Minnesota, 1965–1984. Neurology. 1998 Mar;50(3):735-41. https://doi.org/10.1212/WNL.50.3.735
3. Koul R, Motta A, Kumar A, Razdan S. Status epilepticus in Indian children: A hospital-based study. *J Pediatr Neurosci.* 2014 May-Aug;9(2):131-134. https://doi.org/10.4103/1817-1745.139302
4. Chetan C, Rakesh K, Shetty AK, Jayashree M, Singhi S. Etiological profile and outcome of children with status epilepticus admitted to a tertiary care hospital. *Indian J Pediatr.* 2020 Mar;87(3):194-199. https://doi.org/10.1007/s12098-019-03109-7
5. Sidharth Y, Sharma S, Jain P, Mathur SB, Malhotra RK, Kumar V. Status epilepticus in pediatric patients severity score (STEPSS): A clinical score to predict outcome of status epilepticus in children—A prospec-tive cohort study. *Seizure*. 2019 Oct;71:328-332. https://doi.org/10.1016/j.seizure.2019.09.005

6. Soydan E, Gonullu A, Aksoy Y, Akcay HI, Cengiz M, Ozgurhan G, Duman M. Pediatric Status Epilepticus Severity Score (STEPSS): Predictive performance of functional outcomes. J Child Neurol. 2022 Dec;37(12-14):956-962. https://doi.org/10.1177/08830738221125424
7. Brophy GM, Bell R, Claassen J, Alldredge B, Bleck TP, Glauser T, Laroche SM, Riviello JJ Jr, Shutter L, Sperling MR, Treiman DM, Vespa PM. Guidelines for the evaluation and management of status epilepti-cus. *Neurocrit Care*. 2012 Aug;17(1):3-23. https://doi.org/10.1007/s12028-012-9695-z
8. Yossofzai O, Fallah A, Maniquis C, Ramaswamy V, Snead OC 3rd, Widjaja E, Rutka JT, Ibrahim GM. De-velopment and validation of machine learning models for prediction of seizure freedom after pediatric epilepsy surgery. *Epilepsia*. 2022 Aug;63(8):1956-1969. https://doi.org/10.1111/epi.17320
9. Zhang Y, Yang S, Liu Y, He X, Zhang J, Yu J, Wu Y, Wang Y, Wang D, Wang J, Li Y. Machine learning in predicting mortality of adult status epilepticus: A multicenter validation study. Neurology. 2024 Jan;102(3):241-248. https://doi.org/10.1212/WNL.0000000000002002
10. Loddenkemper T, Goodkin HP, Kapur J. Pediatric status epilepticus—Electroclinical patterns and predic-tors of outcome. Pediatr Neurol. 2021 Feb;117:66-75. https://doi.org/10.1016/j.pediatrneurol.2020.12.014
11. Abend NS, Topjian AA, Ichord RN, Herman ST, Helfaer MA, Donnelly M, Nadkarni V, Clancy RR, Dlugos DJ. Electrographic seizures and status epilepticus in critically ill children. Neurology. 2009 Jun;72(22):1914-1921. https://doi.org/10.1212/WNL.0b013e3181a82687
12. Chin RFM, Verhulst L, Neville BGR, Peters MJ, Scott RC. Treatment of community-onset, childhood convulsive status epilepticus: a prospective, population-based study. *Lancet Neu-rol.* 2008;7(4):335-343.
13. Raspall-Chaure M, Chin RFM, Neville BGR, Scott RC. Outcome of paediatric convulsive status epilepticus: a systematic review. *Lancet Neurol*. 2006;5(9):769-779.

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17365

Polycystic Ovary Syndrome in Medical Students: A Cross-Sectional Study of the Academic and Psychosocial Impact

Ashini A. Patel[1], Caitlyn R. Curley[1], Ganesh Chilukuri[1], Hannah Lee[1], Samantha Snyder[1], Vivek Joshi[2,3]

[1]Doctoral Program, Drexel University College of Medicine, West Reading, Pennsylvania, USA
[2]Department of Biochemistry, Drexel University College of Medicine, West Reading, Pennsylvania, USA
[3]Department of Biochemistry, Kasturba Medical College, Manipal, India

Received: 2025-07-29.
Accepted: 2025-10-05.

J Clin Med Kaz 2025; 22(6): 54–59

Corresponding author:
Vivek Joshi.
E-mail: Vj93@drexel.edu.
ORCID: 0000-0002-0369-9502.

ABSTRACT

Background: Polycystic Ovary Syndrome (PCOS) is a prevalent endocrine disorder affecting up to 13% of reproductive-aged females, often accompanied by physical, metabolic, and psychological challenges. Despite the demanding nature of medical education, limited research explores the burden of PCOS among medical students.

Objective: To evaluate the academic, emotional, and psychosocial impact of PCOS on medical students across four U.S. medical schools.

Methods: A cross-sectional survey was distributed to assigned female at birth medical students aged 18 and older. The questionnaire assessed PCOS diagnosis and treatment, academic performance, symptom burden, emotional well-being, and coping mechanisms. Descriptive statistics and Pearson correlation analyses were performed.

Results: Of the 380 respondents (mean age 25.9), 30.3% reported a PCOS diagnosis—more than double the estimated global prevalence. PCOS was associated with academic stress (34.0%) and reduced work productivity due to emotional distress (70.2%). Symptom burden included menstrual irregularities, weight concerns, bloating, and acne. Emotional impacts were significant: 49.3% reported feeling unattractive, 50.7% worried about infertility, and 50.7% experienced self-blame. While many students demonstrated resilience and adopted positive coping strategies, a notable proportion reported emotional exhaustion and giving up efforts to manage the condition.

Conclusion: Medical students with PCOS experience significant emotional and academic challenges, compounded by the rigors of their training. The prevalence and impact of PCOS in this population highlight the need for targeted institutional support, improved awareness, and mental health resources to promote student well-being and success.

Keywords: PCOS, academic performance, medical students, mental health, quality of life, coping strategies

Introduction

Polycystic Ovary Syndrome (PCOS) is a prevalent endocrine disorder affecting up to 13% of reproductive age females worldwide [1]. Characterized by hyperandrogenism, ovulatory dysfunction, and polycystic ovarian morphology, PCOS presents a complex interplay of metabolic, reproductive, and psychological challenges [1,2]. Despite its significant impact on physical and mental health, research on PCOS remains underfunded and limited, particularly regarding its effects on specific subpopulations. While studies have examined the psychological and social implications of PCOS in the general population, there is a paucity of research investigating its influence on medical students—a group uniquely

positioned at the intersection of academic rigor and healthcare awareness.

Medical students experience heightened stress levels due to demanding coursework, clinical responsibilities, and performance expectations [3]. When coupled with the symptoms of PCOS, such as menstrual irregularities, weight fluctuations, fatigue, and mental health disturbances, their academic performance and overall well-being may be further compromised. Understanding how PCOS affects future physician generations is crucial for identifying tailored interventions that support their educational and professional development.

Emerging evidence suggests a higher prevalence of PCOS among medical students than the general population. One study reported that 15% of 231 medical students surveyed exhibited symptoms consistent with PCOS, indicating a potentially disproportionate burden within this cohort [4]. Moreover, qualitative interviews highlight increased stress, anxiety, and challenges in balancing academic responsibilities among students diagnosed with PCOS. Another study demonstrated that adolescents with PCOS features exhibit metabolic risks, emphasizing the disorder's long-term health implications [5]. However, there is a critical gap in literature exploring the specific experiences of medical students with PCOS and the ways in which the condition influences their academic pursuits.

Our study seeks to address this gap by investigating the experiences of medical students diagnosed with PCOS across four U.S. medical schools. Through a self-administered survey targeting biologically assigned females (AFAB) aged 18 and older, we aim to assess the impact of PCOS on academic performance, mental health, and overall quality of life. Given the established association between PCOS and mental health disorders, our research endeavors to provide data-driven insights that can inform institutional policies and the development of tailored support systems for affected students. By enhancing awareness and fostering institutional accommodations, we hope to improve the academic success and well-being of medical students with PCOS. Research indicates that early intervention in managing PCOS can significantly enhance fertility outcomes. A systematic review and meta-analysis concluded that first-line pharmacological treatments, especially when combining insulin sensitizers with conventional ovulation stimulants, effectively improve clinical pregnancy rates in women with PCOS-related infertility [6]. This underscores the importance of prompt diagnosis and tailored treatment plans for women with PCOS, particularly those navigating the rigorous demands of medical careers.

Methods

Study Design and Participants

This cross-sectional study was conducted between July 2024 and January 2025 across four accredited U.S. medical schools. Participants were recruited via institutional email listservs, student organization announcements, and social media groups. Participation was voluntary, and informed consent was obtained electronically before the start of the survey.

Survey Instrument

Data was collected using a self-administered, anonymous online questionnaire specifically developed for this study. The questionnaire was designed following a review of existing instruments assessing PCOS-related quality of life, coping mechanisms, and academic performance, and was pilot-evaluated on a small sample of medical students (n=10) to ensure clarity and relevance before broader dissemination. The survey included five major sections: demographics (age, ethnicity, academic year, and curriculum type, such as self-directed learning [SDL], lecture-based, or problem-based learning [PBL]); PCOS diagnosis and treatment (questions regarding clinical diagnosis and treatment modalities, including medical, surgical, or alternative therapies); academic performance and stress (assessing perceived academic satisfaction and stress attributed to PCOS); quality of life (measuring the frequency of symptoms such as headaches, bloating, menstrual cramps, weight concerns, excess hair growth, and acne, as well as general health, emotional well-being, physical function, and work productivity); and coping strategies (capturing the extent of engagement in supportive activities, lifestyle adjustments, denial, self-blame, or religious practices). Responses were collected using Likert-type scales appropriate to each section, typically ranging from 1 ("Always") to 5 ("Never") or using similar ordinal formats.

Data Collection

The survey was hosted on a secure online platform (Qualtrics®), and responses were collected anonymously. IP addresses were not stored, ensuring participant confidentiality. Completion of the survey took approximately 10–15 minutes. Data collection continued until a sample size sufficient for meaningful statistical analysis was achieved.

Statistical Analysis

Descriptive statistics, including means, standard deviations, frequencies, and percentages, were used to summarize demographic characteristics, prevalence of PCOS diagnosis, treatment patterns, academic performance, quality of life indicators, and coping strategies.

Bivariate analyses, including Pearson correlation coefficients, were conducted to explore associations between PCOS status, symptom burden, academic stress, coping mechanisms, and emotional well-being. A significance threshold of $p < 0.05$ was used for all inferential statistics.

All analyses were performed using IBM SPSS Statistics version 25.0.

Ethical Considerations

The study was approved by the Institutional Review Boards (IRBs) of all participating institutions. Participation was voluntary, and no identifying information was collected. Participants were informed about the study's objectives, the voluntary nature of participation, and their right to withdraw at any time without any penalty.

Results

Participant Characteristics

A total of 380 participants were included in the study, with a mean age of 25.90 years (SD = 2.66). Among them, 30.3% reported having been diagnosed with PCOS, while 69.7% had not. Of those diagnosed, 86.5% were receiving medical treatment, and 13.5% were using alternative therapies [Table 1].

Academic Performance and Stress

Regarding academic performance, 19.0% of participants reported high satisfaction, 66.2% were satisfied, and 14.8%

Table 1 Demographic and Clinical Characteristics of Participants (N = 380)

Variable	Mean ± SD / Range	Frequency (%)
Age (years)	25.90 ± 2.66	Most common: 25 (18.2%), 27 (17.1%), 28 (18.4%)
Diagnosed with PCOS		30.3%
PCOS Treatment Type		Medical: 86.5% Alternative: 13.5%

Table 2 Academic Performance, Quality of Life, and Emotional Healt

Category	Symptom	Mean ± SD	Most Common Response (%)
Academic Impact	Academic Satisfaction	1.96 ± 0.58	Satisfactory: 66.2%
	Curriculum Type	2.26 ± 1.16	SDL: 38.8%
	Stress due to PCOS	2.06 ± 0.86	No: 39.6%, Yes: 34.0%
Quality of Life	Headaches	4.00 ± 0.85	Sometimes: 43.3%, Never: 31.1%
	Abdominal Bloating	3.49 ± 1.12	Most of the time: 29.8%
	Menstrual Cramps	3.13 ± 1.62	Always: 28.0%
	Irregular Cycles	2.79 ± 1.55	Always: 26.9%, Most: 29.3%
Emotional Health	Concern about Weight	2.15 ± 1.50	Always: 57.8%
	Frustration Losing Weight	2.13 ± 1.48	Always: 53.3%
	Excess Hair Concerns (Q15)	2.73 ± 1.35	Most: 35.1%
	Acne Concerns	2.40 ± 1.59	Always: 45.1%
	Feeling Unattractive	2.30 ± 1.46	Always: 49.3%
	Infertility Concerns	1.49 ± 0.50	Always: 50.7%
	Cancer Concerns	4.01 ± 1.13	Never: 51.2%
General Health	General Health Rating	3.21 ± 0.97	Half the time: 44.9%
	Anxiety About PCOS	3.12 ± 1.24	Half the time: 42.0%

Table 3 Physical Function, Work Productivity, and Coping Strategies

Category	Item / Strategy	Mean ± SD	Most Common Response (%)
Physical Function	Limited in Moderate Activity (Q24)	2.67 ± 0.47	Not at all: 67.3%
	Climb Several Flights of Stairs (Q25)	2.61 ± 0.49	Not at all: 61.5%
Productivity	Less Work Due to Physical Demands (Q26)	1.64 ± 0.48	Yes: 35.9%
	Less Work Due to Emotional Stress (Q27)	1.30 ± 0.46	Yes: 70.2%
Coping	Learning to Live with PCOS (Q38)	3.10 ± 1.00	A lot: 50.7%
	Turning to Activities (Q32)	2.30 ± 1.08	None: 30.9%
	Self-Blame (Q39)	1.92 ± 0.94	A lot: 50.7%
	Use of Substances (Q34)	1.41 ± 0.70	Never: 71.2%
	Seeking Help/Advice (Q35)	2.40 ± 1.05	Medium time: 49.3%
	Gave up Dealing with PCOS (Q36)	1.57 ± 0.74	Yes: 31.7%

reported less satisfaction. Participants followed different curriculum types, including self-directed learning (38.8%), lecture-based (15.8%), and problem-based learning (26.1%). Notably, 34.0% of participants reported experiencing stress about their academic performance being affected by PCOS, while 39.6% did not [Table 2].

Quality of Life and Symptom Burden

PCOS symptoms affected participants' quality of life in multiple ways. The most frequently reported symptoms included headaches (43.3% sometimes), abdominal bloating (29.8% most of the time, 39.3% sometimes), and menstrual cramps (28.0% always, 14.8% half the time). Additionally, 26.9% reported experiencing irregular menstrual cycles always, while 29.3% reported this occurring most of the time. Weight concerns were prevalent, with 57.8% always feeling concerned about being overweight and 53.3% reporting frustration in attempting to lose weight. Excess facial and body hair was also a concern, with 35.1% reporting the need for hair removal most of the time [Table 3].

Emotional and Psychological Impact

Participants expressed significant emotional distress related to PCOS. Approximately 49.3% always felt unattractive due to their weight, body hair, or acne, while 29.0% reported feeling fearful of how others perceived their appearance. Concerns regarding future infertility were reported by 50.7% of participants, while 42.0% felt anxious about their PCOS half the time. Additionally, 51.2% never felt frightened about developing cancer due to PCOS [Table 3].

Physical Function and Work Productivity

Regarding physical activity, 67.3% of participants reported no limitations in moderate activities, while 61.5% were able to climb several flights of stairs without difficulty. Work productivity was affected by physical and emotional demands, with 35.9% reporting accomplishing less work due to physical demands and 70.2% due to emotional stress.

Coping Strategies

Participants adopted various coping strategies to manage PCOS. While 50.7% reported learning to live with PCOS a lot of the time, 66.8% had not turned to alcohol or substances to cope. Instead, 28.2% engaged in activities such as exercise, cooking, or social media as a distraction. Seeking support from others was another strategy, with 49.3% doing so for at least a medium amount of time. However, 31.7% reported having given up trying to deal with PCOS, and 50.7% engaged in self-blame frequently [Table 3].

Discussion

This study offers insights into the experiences of medical students diagnosed with PCOS, shedding light on how it impacts their academic performance, mental health, and overall quality of life. The demographics provide context for how we interpret this data. With a mean age of 25.9 years old, this cohort aligns with the typical age group affected by PCOS. More notably, 30.3% of the participants reported having PCOS, significantly exceeding the global prevalence of approximately 11-13% [1-3]. This suggests a disproportionate burden among this population and emphasizes the need for targeted institutional support.

The symptom burden reported by students diagnosed with PCOS indicates a multifaceted impact on their daily well-being. Common complaints such as abdominal bloating, menstrual cramps, and irregular cycles were reported with high frequency, highlighting the chronic nature of PCOS symptomatology. Furthermore, the was significant concern about weight with 57.8% reporting that they were always concerned and 53.3% reporting frustration with weight loss efforts. This aligns with well-documented challenges related to insulin resistance and metabolic dysfunction in PCOS patients [12-13]. The reports of physical discomfort and body image concerns related to hirsutism, acne, and bloating further negatively impact quality of life and contribute to emotional distress. Similar findings showed that women with PCOS have impaired health-related quality of life, particularly driven by symptoms related to physical appearance and menstrual irregularities, which frequently contribute to psychological distress [7]. The stress burden can impact their academic performance. In our study, 34% of respondents reported experiencing stress specifically related to concerns about how PCOS was impacting their academic performance, which has been corroborated with previous studies highlighting the increased academic and emotional strain experienced by medical students with PCOS [12-14].

Emotional distress was also highly prevalent among participants diagnosed with PCOS. Nearly half (49.3%) reported persistent feelings of unattractiveness related to symptoms like weight issues, acne, or hirsutism, which are manifestations often associated with hyperandrogenism [3,5]. Additionally, 50.7% expressed anxiety about future infertility, highlighting the specific reproductive concerns within this group. PCOS is one of the leading causes of female infertility, as irregular ovulation or anovulation can significantly reduce the chances of conception [15]. This can be particularly distressing for female physicians, who often face delays in family planning due to the demands of medical training and professional responsibilities. Given that many female physicians postpone family planning due to career demands [9], the interplay between professional goals and reproductive health challenges in women with PCOS deserves particular attention. Prior research supports that infertility-related stress is notably elevated in women with PCOS, further amplifying anxiety in an already high-pressure environment like medical training [10,14].

Regarding physical function and work productivity, more than two-thirds indicated no limitations with moderate activity related to their PCOS. However, 70.2% of participants reported accomplishing less work due to emotional demands. Huddleston et al. (2024) reported similar findings, with 51.5% of participants stating that PCOS interfered with their work performance, primarily due to symptoms of anxiety and depression. The disparity between physical capability and reduced productivity highlights how emotional health plays a critical role in academic and professional performance [14].

Finally, students reported a variety of ways they coped in response to their PCOS symptoms. Of note, just over half (50.7%) reported learning to manage and accept the condition over time. Most students did not turn to maladaptive behaviors like substance use, and many cited positive outlets such as hobbies or social activities like spending time with friends. These strategies reflect a sense of resilience and the ability to adapt despite ongoing challenges. However, these findings also showed that 31.7% of students said they had given up trying to manage their PCOS and 50.7% admitted blaming themselves for their symptoms [8-10]. These responses point to internalization of the stigma around PCOS and also to feelings of burnout that come from managing chronic conditions, while also managing the rigorous demands of medical training. Overall, this highlights a potential gap in accessible mental health resources tailored to students dealing with long-term health issues.

Given the high prevalence and multifaceted burden of PCOS among medical students, medical institutions must consider implementing targeted support mechanisms. These could include improved access to on-campus mental health services with providers knowledgeable about chronic health conditions like PCOS. Academic accommodations such as flexible scheduling, extensions on coursework, or medical leave policies should be clearly outlined and accessible to affected students. Institutions could also consider offering peer-support groups or wellness workshops specifically addressing reproductive and hormonal health13. Finally, training faculty and academic advisors to recognize and respond to the unique challenges faced by students with PCOS may foster a more inclusive educational environment [15]. Integrating such resources within medical education infrastructure can help mitigate the emotional and academic toll of PCOS, ultimately promoting student retention, well-being, and professional success.

Further studies with more diverse samples and longitudinal follow-up would be beneficial to build on the findings.

Conclusion

This study underscores the complex burden PCOS places on medical students, highlighting a prevalence notably higher than that seen in the general population. Beyond the physical

symptoms, many students reported significant emotional strain, including anxiety, concerns about fertility, and struggles with body image. Although most participants continued to function well physically and expressed general academic satisfaction, the emotional weight of managing a chronic condition alongside the pressures of medical training was evident. Some students demonstrated strong coping skills and resilience, but others described emotional exhaustion, self-blame, and a sense of discouragement. These findings point to a clear need for medical schools to offer more targeted and accessible support systems for students managing PCOS.

Limitations

The main limitation to this study is, the study relies on self-reported data, which introduces the potential for recall bias and social desirability bias, especially regarding sensitive topics such as body image and infertility concerns. The other limitation is, the cross-sectional design precludes any temporal or causal interpretations between PCOS and academic or emotional outcomes. Lastly, there is a risk of selection bias, as students with more severe symptoms may have been more motivated to respond, possibly inflating prevalence and impact estimates.

Author Contributions: Conceptualization, A.P., V.J., G.C., C.C., H.L., and S.S.; methodology, A.P., V.J., G.C., C.C., H.L., and S.S.; formal analysis, A.P., V.J., G.C., C.C., H.L., and S.S.; investigation, A.P., V.J., G.C., C.C., H.L., and S.S.; resources, A.P., V.J., G.C., C.C., H.L., and S.S.; data curation, A.P., V.J., G.C., C.C., H.L., and S.S.; writing – original draft preparation, V.J., C.C., G.C., S.S., H.L., and A.P.; writing – review and editing, A.P., V.J., G.C., C.C., H.L., and S.S.; visualization, A.P., V.J., G.C., C.C., H.L., and S.S.; supervision, A.P., V.J., G.C., C.C., H.L., and S.S. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The data generated during the research and analysis are not available publicly but are available from the corresponding author on a reasonable request.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. Stener VE., Teede H, Norman RJ, Legro RS, Goodarzi, MO, Dokras A, Piltonen, TT. Polycystic ovary syn-drome. *Nature Reviews Disease Primers.* 2024; 10(1): 27. https://doi.org/10.1038/s41572-024-00484-9
2. Azziz R, Carmina E, Chen Z, Dunaif A, Laven JS, Legro RS, Lizneva D. Polycystic ovary syndrome. *Nature Reviews Disease Primers*. 2016; 2: 16057. https://doi.org/10.1038/nrdp.2016.57
3. Dyrbye LN, Thomas MR, Shanafelt TD. Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. Academic Medicine. 2006; 81(4): 354–373. https://doi.org/10.1097/00001888-200604000-00009
4. Aziz R, Sthanadar AA, Liaqat A, Noor S. Prevalence of polycystic ovary syndrome among medical stu-dents. *Journal of Ayub Medical College Abbottabad.* 2005; 17(4): 72–75.
5. Hart R, Hickey M, Franks S. Definitions, prevalence and symptoms of polycystic ovaries and polycystic ovary syndrome. Best Practice & Research Clinical Obstetrics & Gynaecology. 2004; 18(5): 671–683. https://doi.org/10.1016/j.bpobgyn.2004.05.006
6. PMC. Impact of early intervention on fertility outcomes in PCOS: A systematic review. *PubMed Central.* 2025: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9983155/
7. Li Y, Ng EH, Stener-Victorin EH, Wu LT, Wu X. Polycystic ovary syndrome is associated with negatively variable impacts on domains of health-related quality of life: Evidence from a meta-analysis. *Fertility and Sterility.* 2011; 96(2): 452–458. https://doi.org/10.1016/j.fertnstert.2011.05.090
8. Aspire Fertility. Polycystic ovarian syndrome and fertility. *Aspire Fertility.* 2025: https://www.aspirefertility.com/fertility-treatment/polycystic-ovarian-syndrome
9. Medscape. More women entering medical school than men for the sixth straight year. *Medscape*. 2025: https://www.medscape.com/viewarticle/more-women-entering-medical-school-than-men-sixth-straight-2025a100021y
10. Dokras A, Clifton S, Futterweit W, Wild R. Increased prevalence of anxiety symptoms in women with pol-ycystic ovary syndrome: Systematic review and meta-analysis. *Fertility and Sterility.* 2012; 97(1): 225–230. https://doi.org/10.1016/j.fertnstert.2011.10.022
11. Huddleston HG, Milani A, Blank R. Productivity loss due to polycystic ovary syndrome and its relation-ship to race, mental health, and healthcare delivery indices. *F&S Reports.* 2024; 5(2): 157–163. https://doi.org/10.1016/j.xfgr.2023.12.004.
12. Kupoluyi JA. Intimate partner violence as a factor in contraceptive discontinuation among sexually active married women in Nigeria. *BMC Women's Health.* 2020; 128. https://doi.org/10.1186/s12905-020-00990-y.
13. Murugesu S, Saso S, Jones BP, Bracewell-Milnes T, Athanasiou T, Mania A, Serhal P, Ben-Nagi J. Does the use of calcium ionophore during artificial oocyte activation demonstrate an effect on pregnancy rate? A meta-analysis. *Fertil Steril.* 2017; 108(3):468-482. https://doi.org/10.1016/j.fertnstert.2017.06.029. PMID: 28865547.

14. Murri M, Luque-Ramírez M, Insenser M, Ojeda-Ojeda M, Escobar-Morreale HF. Circulating markers of oxidative stress and polycystic ovary syndrome (PCOS): a systematic review and meta-analysis. *Hum Re-prod Update.* 2013;19(3):268-88. https://doi.org/10.1093/humupd/dms059.
15. Ofei-Dodoo S, Moser SE, Kellerman R, Wipperman J, Paolo A. Burnout and Other Types of Emotional Distress Among Medical Students. Med Sci Educ. 2019; 29(4):1061-1069. https://doi.org/10.1007/s40670-019-00810-5.

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17366

The Impact of Glucagon-like Peptide-1 Receptor Agonists on Obese Individuals' Quality of Life and Eating Behaviors

Aya S. Aldewachi[1], Mohammed I. Aladul[1]

[1]Department of Clinical Pharmacy, College of Pharmacy, University of Mosul, Ninevah, Iraq

Received: 2025-06-30.
Accepted: 2025-09-28.

J Clin Med Kaz 2025; 22(6): 60–72

Corresponding author:
Mohammed I. Aladul.
Email: m.i.m.aladul@uomosul.edu.iq.
ORCID: 0000-0002-7368-3469.

ABSTRACT

Background: Obesity is a multifactorial, neurobehavioral, chronic disease associated with long-term metabolic and psychosocial health consequences, all of which affect the quality of life (QoL). Eating behavior is an essential and critical factor that is connected to both the development of obesity and its effect on QoL. While glucagon-like peptide-1 receptor agonists (GLP-1RAs) show promise for improving both weight and metabolic parameters, their effects on specific eating behaviors, particularly emotional eating, remain incompletely understood.

Objectives: This study aims to investigate the complex relationships between obesity, eating behaviors, and QoL in individuals with different BMI categories.

Methods: This cross-sectional study utilized both online and face-to-face surveys conducted among a cohort of Iraqi individuals with BMI higher than 18.5 which were users and non-users of GLP-1RAs between December 2024 and March 2025.

Results: Participants with higher BMI demonstrated higher scores of emotional overeating (EOE) and hunger, while satiety responsiveness and slowness in eating were inversely related. Hunger scores were lower in GLP-1RAs users with higher QoL scores, though EOE remained elevated. Behavioral clustering demonstrated that the fuzzy eaters' group are associated with high EOE, and low satiety, while the gourmets' group are associated with higher food enjoyment and slower eating.

Conclusion: Despite improving appetite regulation and QoL, GLP-1RAs do not completely address emotional eating. Behavioral clustering provides an insight to understand individual differences in obesity-related behavior and necessitates the need for personalized interventions. Physiologic and psychological dimensions of obesity must both be dealt with for the improvement of patient well-being.

Keywords: Obesity, eating behavior, quality of life, semaglutide, liraglutide.

Introduction

The World Obesity Federation defines obesity as "a multifactorial, neurobehavioral, chronic progressive, relapsing, and treatable disease, leading to adverse metabolic, biomechanical, and psychosocial health consequences" [1]. According to the World Health Organization (WHO), in 2022, 2.5 billion adults aged 18 years and above were overweight, and about 16% of adults were obese in the same year. The prevalence of obesity has more than doubled between 1990 and 2022 [2]. In Iraq, where this study was conducted, obesity prevalence exceeds regional averages at 26.5% in men and 40.1% in women [3], with recent studies identifying distinct eating behavior patterns like emotional overeating (EOE) as key drivers of weight gain in this population [4].

Obesity is associated with long-term health consequences, including type 2 diabetes, hypertension,

cardiovascular diseases, respiratory complications, certain cancers, reproductive health issues, and osteoarthritis [5]. The occurrence of these health problems increases with body mass index (BMI), regardless of age and gender [2]. In addition to physical problems, obesity can lead to neuropsychiatric conditions, including anxiety, poor self-image, sleep disorders, and depression, all of which affect the quality of life (QoL) of individuals with obesity [6].

Eating behavior is another essential and critical factor that is connected to both the development of obesity and its effect on QoL. EOE, binge eating, and poor satiety responsiveness are types of eating behaviors that are prevalent among individuals with obesity or excess body weight and are linked with increased caloric intake and weak dietary regulation [7]. These traits, particularly the "fuzzy eater" profile characterized by EOE and poor satiety response, recently identified in Iraqi adults [4], may contribute to weight gain and hinder weight management, thereby contributing to emotional distress and physical health decline [8]. In this study, the terms 'eating behavior' and 'appetitive trait' are used interchangeably to describe the cognitive, emotional, and physiological factors that influence how individuals interact with food. While appetitive traits often reflect stable, biologically influenced tendencies (e.g., satiety responsiveness), eating behaviors encompass both habitual and situational expressions of these traits.

Pharmacological interventions for obesity have emerged as a complementary approach to lifestyle changes. Glucagon-like peptide-1 receptor agonists (GLP-1RAs) have demonstrated promising results in weight reduction. While GLP-1RAs show promise for improving both weight and metabolic parameters [9, 10], their effects on specific eating behaviors, particularly EOE, remain incompletely understood. Notably, Charlotte C. van Ruiten et al. [11] study found that EOE blunts the appetite-suppressing effects of GLP-1RAs, suggesting these behaviors may persist despite treatment. This gap is especially relevant in Iraq, where cultural dietary practices and prevalent dysregulated eating patterns [4] could further modulate treatment response. Furthermore, the potential QoL benefits beyond weight reduction warrant investigation [11].

This study aims to examine the complex relationships between obesity, eating behaviors, and QoL in Iraqi adults across different BMI categories. We aim to understand how increasing BMI is associated with changes in QoL and shifts in eating behaviors, such as EOE and satiety responsiveness, with a focus on the "fuzzy eater" profiles observed in this population [12]. By comparing GLP-1RA users and non-users, we will assess the potential impact of these medications on appetite regulation and satiety, testing the hypothesis—supported by Charlotte C. van Ruiten et al. [11]—that EOE might influence the effectiveness of treatment. Additionally, we will explore whether GLP-1RAs provide psychosocial benefits beyond physical health improvements, filling an important gap in obesity management strategies within the Iraqi context. Although the study is cross-sectional in nature, these insights could inform future longitudinal research into the long-term effects of GLP-1RA treatment.

Methods

Study Design

This cross-sectional study assessed QoL and eating behaviors through both online and face-to-face surveys conducted among overweight and obese Iraqi adults, including users and non-users of GLP-1RAs.

Participants

The study included Iraqi adults with a BMI higher than 18.5, who were either users or non-users of GLP-1RAs. Participants were selected using a convenience sampling approach, with recruitment conducted in dietetics clinics, healthcare centers, and community pharmacies across various locations. Eligible participants were required to meet the following criteria: (1) age between 18 and 55 years, (2) BMI greater than 18.5, (3) availability of informed consent, and (4) ability to understand and complete the survey in Arabic. Exclusion criteria included (1) individuals with severe comorbidities (e.g., type 1 diabetes, severe psychiatric conditions) or (2) those not willing to participate in the study.

Follow-up

As a cross-sectional study, there was no follow-up required. Data collection was conducted between December 2024 and March 2025, with participants completing surveys during a one-time visit to participating healthcare centers or via online platforms.

Matching

This study did not involve matched groups. However, GLP-1RA users and non-users were compared across key outcomes. The two groups were not matched based on demographic factors (e.g., age, gender), but statistical controls were used in the analysis to adjust for potential confounders such as age, gender, and BMI.

Variables

- Outcomes: The primary outcomes included EOE, hunger, satiety responsiveness, and overall QoL, measured using the Arabic version of the Adult Eating Behavior Questionnaire (AEBQ) and the Quality of Life, Obesity, and Dietetics (QOLOD) scale.
- Exposures: The exposure of interest was the use of GLP-1RAs (liraglutide, semaglutide, or tirzepatide).
- Predictors: BMI category (normal weight, overweight, obese class I–III), gender, age, marital status, educational level, and occupation were considered as potential predictors of eating behaviors and QoL outcomes.
- Confounders: Potential confounders included age, gender, BMI, and use of other anti-obesity medications (e.g., Orlistat or herbal products).
- Effect Modifiers: Gender and educational level were examined as potential effect modifiers, as these factors may influence eating behaviors and QoL differently.

Data Sources/Measurement

- Data sources: Data were collected through surveys administered online via Microsoft Forms and in-person at healthcare facilities. The AEBQ was used to assess eating behaviors such as EOE, food enjoyment, and hunger. The QOLOD scale was used to measure QoL.
- Assessment Methods: The AEBQ is a validated 27-item questionnaire assessing six subscales (hunger, EOE, food enjoyment, satiety responsiveness, food fussiness, and slowness in eating). QoL was assessed using the QOLOD scale, which evaluates physical impact, psychosocial impact, and diet experience in obese populations.
- Comparability: The methods of assessment (AEBQ and QOLOD) were consistently applied to both GLP-1RA users and non-users to ensure comparability of the data.

Bias

Efforts to minimize bias included:

1. Randomization of participants to the online or in-person survey method to reduce selection bias.

2. Use of validated and reliable instruments (AEBQ and QOLOD) to minimize measurement bias.

3. Anonymization of survey responses to reduce social desirability bias.

However, the potential for recall bias exists, particularly with self-reported measures of eating behavior and QoL.

Sample Size

The sample size was calculated to ensure sufficient power for detecting differences in eating behaviors and QoL between GLP-1RA users and non-users. Based on prior studies, an expected moderate effect size (Cohen's d = 0.5) and a significance level of 0.05, a minimum of 300 participants per group (users and non-users) was needed to achieve 80% power. A total of 994 participants were included in the study, which provided ample statistical power for the analysis.

Recruitment

Recruitment employed a dual approach: trained researchers conducted in-person interviews at healthcare facilities, while electronic participation was facilitated through QR codes posted in these locations and distributed via social media platforms, including WhatsApp and closed Facebook groups.

Participants were categorized into two distinct groups for comparison: non-users of GLP-1RAs and active users who had been taking specific medications (liraglutide/Saxenda®, semaglutide/Ozempic®, or tirzepatide/Mounjaro®) for weight management for at least four weeks prior to participation. The sampling strategy intentionally sought diversity across demographic factors. All participants provided both written and verbal informed consent after receiving complete study information, with no incentives offered for participation.

Instruments

The comprehensive survey instrument incorporated three key domains. The first section gathered essential demographic characteristics, including but not limited to age, weight, height, and other relevant variables. For the critical behavioral assessment, the research team administered the validated Arabic version of the AEBQ [12], which has demonstrated reliability in Arabic-speaking populations. This 27-item instrument thoroughly evaluates six core subscales through a 5-point Likert scale ranging from "strongly agree" to "strongly disagree": hunger, EOE, food enjoyment, satiety responsiveness, food fussiness, and slowness in eating.

QoL measurement utilized the Arabic version of the QOLOD rating scale [13], previously validated specifically for obese populations in Iraq. This sensitive 25-item tool captures three crucial dimensions: physical impact, psychosocial impact, and diet experience.

The research team implemented rigorous privacy protections throughout the study—questionnaires contained no personally identifiable questions, all responses were anonymized immediately upon collection, assigned unique codes, and stored securely with restricted access to ensure complete confidentiality.

Data Collection

For electronic data collection, the team used the Microsoft Forms platform to host the survey. The digital version was distributed through multiple channels: direct links shared in targeted social media groups and QR codes physically posted in participating healthcare locations. Several collaborating dietitians assisted with recruitment by distributing informational cards containing the survey QR code to their eligible patients. Face-to-face administration followed standardized protocols, with trained researchers conducting interviews at selected healthcare sites. Participants typically completed the survey within 10-15 minutes. The data collection period extended from December 1, 2024, to March 30, 2025, to ensure adequate sample representation across seasons.

Statistical Analysis

The response data were retrieved and analyzed by the authors on April 6, 2025. The statistical analysis employed R software version 4.3 for all computations. Initial descriptive analyses included the calculation of means, standard deviations, medians, interquartile ranges (IQR), frequencies, and percentages to characterize both demographic variables and scores on the eating behavior scales. The research team used the Elbow method to objectively determine the optimal number of clusters for analyzing the six-food approach and avoidance traits, which clearly indicated a two-cluster solution. They then performed K-means cluster analysis to divide participants into homogeneous behavioral groups based on their patterns of scores across all appetite-related traits.

Given that Shapiro-Wilk tests confirmed non-normal distribution for most continuous outcome variables (including total scores from both the AEBQ and QOLOD subscales), the team appropriately selected non-parametric statistical methods for inferential analyses. The Mann-Whitney U test compared differences between two independent groups, while the Kruskal-Wallis test extended this to multiple group comparisons. To examine relationships between variables, the analysis incorporated both Spearman's correlation coefficients and multiple linear regression models that controlled for potential confounding factors. Throughout all analyses, the team maintained a consistent significance threshold of $p < 0.05$ for determining statistically meaningful results.

Ethical Considerations

The study protocol received full approval from the Pharmaceutical Research Ethics Committee at the College of Pharmacy, University of Mosul, Iraq, before the commencement of data collection (PREC-24-25-4-1 in 2nd Sep 2024).

Results

The study included a total of 994 participants, with a median age of 31 years (interquartile range [IQR] = 27–35), ranging from 18 to 55 years. The majority of participants were female (845; 85%), married (792; 79.7%), employed (561; 56.4%), and had attained university-level education (686; 69%). Median BMI was 28.65 kg/m² (IQR = 25.0–35.0), median height was 163 cm (IQR = 158–170), and median weight was 75.25 kg (IQR = 63–88). Participants were classified into five BMI categories: normal weight (n = 239; 24%), overweight (n = 340; 34.2%), obese class I (n = 223; 22.4%), obese class II (n = 122; 12.3%), and obese class III (n = 70; 7%). Of these, 257 (25.9%) were current users of GLP-1RAs, with Ozempic® (n = 107; 41.6%), Saxenda® (n = 99; 38.5%), and Mounjaro® (n = 51; 19.8%) being the most frequently used agents (Tables 1 and 2).

Table 1 Categorical demographic characteristics of study participants

Variable		Frequency (%)
Gender	Male	149 (15%)
	Female	845 (85%)
Obesity category	Normal weight (18.5-24.9)	239 (24%)
	Overweight (25.0-29.9)	340 (34.2%)
	Obese class I (30.0-34.9)	223 (22.4%)
	Obese class II (35.0-39.9)	122 (12.3%)
	Obese class III ≥ 40	70 (7%)
Marital status	Single	202 (20.3%)
	Married	792 (79.7%)
Occupation	Student	134 (13.5%)
	Employer	561 (56.4%)
	Not working	299 (30.1%)
Educational level	Primary education	20 (2%)
	Secondary education	120 (12.1%)
	University education	686 (69%)
	Postgraduate education	168 (16.9%)
Obesity drug used	No drug	631 (63.5%)
	Orlistat	40 (4.0%)
	Herbals	66 (6.6%)
	GLP-1RAs	257 (25.9%)
Type of GLP-1RAs	Ozempic®	107 (41.6%)
	Saxenda®	99 (38.5%)
	Mounjaro®	51 (19.8%)

GLP-1RA = Glucagon-like peptide-1 receptor agonists, BMI = Body mass index. Data are presented as Data are presented as frequency (% percentages) for categorical variables. Descriptive statistics are used to summarize the demographic characteristics of the study participants.

Table 2 Continuous demographic characteristics of study participants

Variable	Mean	SD	Median	IQR (Q1–Q3)	Min	Max	SHAPIRO-WILK P-value
Age	31.4	6.2	31.0	27.0–35.0	18.0	55.0	< 0.001
BMI	29.73	6.14	28.65	25.0–35.0	20.0	56.75	0.021
Height (cm)	164.0	8.5	163.0	158.0–170.0	145.0	190.0	0.035
Weight (kg)	78.85	18.48	75.25	63.0–88.0	45.0	164.0	0.041

BMI = Body Mass Index, SD = Standard Deviation, IQR = Interquartile Range, Min = Minimum, Max = Maximum. Data are presented as mean ± standard deviation (SD), median, and interquartile range (IQR) for continuous variables. The Shapiro-Wilk test was used to assess the normality of the distribution for each variable. A p-value of < 0.05 indicates that the variable does not follow a normal distribution.

Appetitive Traits and Quality of Life Measures

Appetitive traits were assessed using the Arabic version of the Adult Eating Behavior Questionnaire (AEBQ). All AEBQ subscales deviated significantly from normality (all p-values < 0.05), and are thus presented as medians and IQRs. EOE had a median score of 14 (IQR = 12–18), food fussiness was 12 (IQR = 11–14), hunger was 12 (IQR = 10–14), slowness in eating was 11 (IQR = 9–13), satiety responsiveness was 12 (IQR = 9–13), and enjoyment of food was 20 (IQR = 18–23) (Table 3).

QoL was measured using the QOLOD rating scale. The total QOLOD score was the only variable found to be normally distributed (Shapiro-Wilk p = 0.358) and is reported as mean ± standard deviation (SD): 67.93 ± 17.85. Physical impact had a median score of 34 (IQR = 28–36), psycho-social impact was 29 (IQR = 22–33), and diet experience was 13 (IQR = 11–16). These

Table 3 Appetitive traits score using adult eating behavior questionnaire (AEBQ) and quality of life measures using QOLOD

Variable	Mean	SD	Median	IQR (Q1–Q3)	Min	Max	SHAPIRO-WILK SIG. (P)
Appetitive traits							
Enjoyment of food	20.57	3.26	20.00	[18–23]	10.0	28.0	0.0112
Emotional overeating	14.71	5.79	14.00	[12–18]	5.0	25.0	0.000
Food fussiness	12.29	2.26	12.00	[11–14]	5.0	20.0	0.000
Hunger	12.03	3.05	12.00	[10–14]	4.0	20.0	0.000
Slowness in eating	11.07	3.65	11.00	[9–13]	4.0	20.0	0.000
Satiety responsiveness	12.51	3.23	12.00	[9–13]	4.0	20.0	0.000
Quality of life							
Physical impact	34.07	9.26	34.00	[28–36]	11.0	55.0	0.000
Psycho-social impact	29.55	8.75	29.00	[22–33]	9.0	45.0	0.000
Diet experience	14.59	4.45	13.00	[11–16]	5.0	25.0	0.000
Total QOLOD score	67.93	17.85	79.5	[65-92]	25.0	113.0	0.358

AEBQ = Adult Eating Behavior Questionnaire, QOLOD = Quality of Life, Obesity and Dietetics, SD = Standard Deviation, IQR = Interquartile Range, Min = Minimum, Max = Maximum. Data are presented as mean ± standard deviation (SD), median, and interquartile range (IQR) for continuous variables. The Shapiro-Wilk test was used to assess the normality of the distribution for each variable. A p-value of < 0.05 indicates that the variable does not follow a normal distribution.

scores indicate that participants experienced moderate-to-severe impairment in health-related QoL due to obesity (Table 3).

Difference in appetitive traits scores and quality-of-life subscales scores between GLP-1RAs users and non-users

Significant differences were observed between GLP-1RA users and non-users in certain appetitive traits. EOE scores were significantly higher among GLP-1RA users compared to non-users (median = 16 vs. 14, p = 0.001). Hunger levels were lower in GLP-1RA users (median = 11.69 vs. 12.14, p = 0.022). No statistically significant differences were found between the two groups in terms of enjoyment of food, food fussiness, slowness in eating, or satiety responsiveness.

QoL measures showed consistent and statistically significant differences between GLP-1RA users and non-users across all domains. Non-users reported greater physical impact (median = 35.53 vs. 29.86, p = 0.001), higher psycho-social impact (median = 30.83 vs. 25.83, p = 0.001), worse diet experience (median = 15.41 vs. 12.23, p = 0.001), and overall poorer QoL (mean total QOLOD score = 81.79 vs. 67.93, p = 0.001). These findings suggest that GLP-1RA use is associated with improved health-related QoL outcomes in individuals across the weight spectrum (Table 4).

Table 4 Difference in appetitive traits scores and quality-of-life subscales scores between GLP-1RAs users and non-users

Variable	GLP-1RAs non users Mean (SD)	GLP-1RAs users Mean (SD)	P-value
Enjoyment of food	20.65 (2.95)	20.29 (3.99)	0.479
Emotional overeating	14.19 (5.71)	16.21 (5.74)	0.001
Food fussiness	12.26 (2.19)	12.37 (2.43)	0.6
Hunger	12.14 (2.88)	11.688 (3.47)	0.022
Slowness in eating	11.10 (3.59)	10.97 (3.80)	0.52
Satiety responsiveness	12.35 (2.90)	12.92 (3.99)	0.114
Physical impact	35.53 (8.70)	29.86 (9.51)	0.001
Psycho-social impact	30.83 (8.41)	25.83 (8.41)	0.001
Diet experience	15.41 (4.33)	12.23 (3.93)	0.001
Total QOLOD	81.788 (18.37)	67.93 (17.85)	0.001

GLP-1RA = Glucagon-like peptide-1 receptor agonists, SD = Standard Deviation, QOLOD = Quality of Life, Obesity and Dietetics. Data are presented as mean ± standard deviation (SD). Mann-Whitney U tests were used to compare the differences between GLP-1RA users and non-users. Statistical significance was set at p < 0.05.

Differences in Appetitive Traits Across Demographic Groups

Significant differences were observed in several appetitive traits across different demographic categories. Gender was associated with variations in enjoyment of food, EOE, hunger, and slowness in eating. Males reported significantly lower enjoyment of food (17.36 vs. 21.03, p = 0.001) and exhibited reduced EOE (12.28 vs. 17.21, p = 0.001) compared to females. Hunger levels were also significantly lower among males (9.63 vs. 12.20, p = 0.001), and slowness in eating showed a trend toward being lower in males, though this did not reach strong statistical significance (12.05 vs. 10.69, p = 0.049). No significant gender-based differences were found in food fussiness or satiety responsiveness.

When stratifying by BMI category, EOE was significantly higher in individuals classified as obese class I (17.89) compared to those who were normal weight (15.81, p = 0.001). Satiety responsiveness decreased with increasing BMI, with the lowest scores observed in overweight individuals (14.72) and the highest in obese class III participants (13.86, p = 0.001). Enjoyment of food did not show a consistent pattern across BMI groups (p = 0.160), nor did food fussiness (p = 0.066). However, hunger appeared to decline slightly in more severe obesity categories, particularly in obese class III (10.57, p = 0.088).

Marital status did not yield any statistically significant differences in appetitive traits after correction for multiple comparisons, although single individuals tended to report slightly higher enjoyment of food (21.54 vs. 20.11, p = 0.077) than married participants.

Occupational status was linked to differences in enjoyment of food, EOE, and satiety responsiveness. Students scored significantly lower in EOE (14.31 vs. 16.50, p = 0.022) and enjoyment of food (19.25 vs. 20.67, p = 0.003) compared to employers. Additionally, students had lower satiety responsiveness (13.31 vs. 12.79, p = 0.012), suggesting a potential link between employment status and appetite regulation behaviors.

Educational level showed significant associations with multiple appetitive traits. Individuals with primary education reported lower enjoyment of food (19.83) compared to postgraduate-educated participants (21.45, p = 0.001), and EOE was also lower in this group (1
6
16 vs. 18.35, p = 0.002). Higher educational attainment was associated with increased satiety responsiveness, with postgraduate participants showing the highest scores (11.29) and secondary-educated individuals the lowest (15.02, p = 0.001).

Regarding anti-obesity medication use, EOE was significantly higher among GLP-1RA users (16.21) compared to non-users (13.75, p = 0.001). Users of herbal products also showed elevated EOE (17.59) compared to other medication groups. Hunger was significantly lower in non-users (12.1) than in users of herbal products (12.65, p = 0.014). No major differences were found in satiety responsiveness across medication groups except for a marginally lower score in non-users (12.49 vs. 12.92, p = 0.027).

Finally, analysis by type of GLP-1RA used revealed that Ozempic® users had significantly higher enjoyment of food (21.17) than Saxenda® users (19.54, p = 0.020). EOE was also significantly higher in Ozempic® users (17.11) compared to Saxenda® users (15.16, p = 0.042). These findings suggest subtle but meaningful behavioral distinctions between GLP-1RA types, warranting further investigation into their differential effects on eating behavior profiles (Table 5).

Differences in Quality of Life Across Demographic Groups

QoL was assessed using the QOLOD rating scale, which includes physical impact, psycho-social impact, diet experience, and total QoL scores. Significant differences in QoL

Table 5 Difference in Appetitive traits scores between each demographic group

Demographic variable		EF		EOE		FF		H		SE		SR	
		Mean (SD)	P value	Mean (SD)	P value	Mean (SD)	P value	Mean (SD)	P value	Mean (SD)	P value	Mean (SD)	P value
Gender	Male	17.36 (4.80)	0.001	12.28 (4.92)	0.001	13.13 (2.65)	0.171	9.63 (4.15)	0.001	12.05 (5.04)	0.049	15.67 (5.08	0.582
	Female	21.03 (3.38)		17.21 (5.51)		12.18 (2.34)		12.20 (3.07)		10.69 (3.38)		12.22 (3.33)	
Obesity category	Normal weight	20.18 (3.40)	0.160	15.81 (6.32)	0.001	13.00 (2.82)	0.066	12.18 (2.52)	0.088	10.63 (3.13)	0.001		0.001
	Overweight	20.78 (3.44)		15.83 (5.62)		12.34 (2.25)		12.24 (2.86)		10.71 (3.62)		14.72 (2.28)	
	Obese class I	21.06 (3.49)		17.89 (5.18)		12.09 (2.40)		12.10 (3.03)		10.40 (3.24)		12.24 (3.17)	
	Obese class II	19.87 (3.83)		15.89 (5.26)		12.22 (2.49)		11.12 (3.46)		11.96 (4.25)		11.81 (3.38)	
	Obese class III	18.60 (5.50)		14.35 (6.87)		13.00 (2.59)		10.57 (4.95)		11.17 (4.43)		13.86 (4.34)	
Marital status	Single	21.54 (3.33)	0.077	17.51 (5.71)	0.156	12.93 (1.96)	0.124	12.48 (3.03)	0.243	10.09 (2.90)	0.167	13.03 (2.87)	0.648
	Married	20.11 (4.05)		16.02 (5.74)		12.29 (2.48)		11.57 (3.52)		11.10 (3.91)		12.91 (4.13)	
Occu-pation	Student	19.25 (4.39)	0.003	14.31 (6.19)	0.022	12.56 (1.99)	0.702	11.21 (4.29)	0.113	10.50 (4.29)	0.132	13.31 (4.53)	0.012
	Employer	20.67 (3.94)		16.50 (5.74)		12.34 (2.50)		11.68 (3.40)		10.91 (3.87)		12.79 (3.98)	
	Not working	19.74 (3.78)		16.43 (5.39)		12.34 (2.48)		11.98 (3.16)		11.41 (3.28)		13.10 (3.70)	
Educa-tional level	Primary education	19.83 (2.22)	0.001	16.16 (5.81)	0.002	12.83 (1.60)	0.524	13.16 (3.12)	0.834	12.66 (2.80)	0.013	12.00 (3.28)	0.001
	Secondary education	17.63 (4.34)		13.44 (6.49)		13.08 (2.33)		10.33 (4.15)		12.61 (4.03)		15.02 (4.58)	
	University education	20.50 (3.94)		16.08 (5.43)		12.32 (2.50)		11.68 (3.40)		11.04 (3.72)		13.06 (3.77)	
	Postgraduate education	21.45 (3.31)		18.35 (5.40)		12.01 (2.32)		12.38 (3.01)		9.56 (3.51)		11.29 (3.62)	
Obesity drug used	No drug	20.61 (2.97)	0.323	13.75 (5.55)	0.001	12.31 (2.19)	0.484	12.1 (2.90)	0.014	11.14 (3.60)	0.775	12.49 (2.89)	0.027
	Orlistat	20.37 (2.52)		15.50 (6.10)		12.30 (2.24)		11.37 (2.89)		11.12 (3.91)		11.40 (3.11)	
	Herbals	21.21 (2.94)		17.59 (5.83)		11.74 (2.13)		12.65 (2.51)		10.74 (3.27)		11.68 (2.73)	
	GLP-1RAs	20.29 (3.99)		16.21 (5.74)		12.37 (2.43)		11.68 (3.47)		10.97 (3.80)		12.92 (3.99)	
GLP-1RAS using or not	Ozempic®	21.17 (3.73)	0.020	17.11 (5.83)	0.042	12.09 (2.43)	0.264	12.37 (3.24)	0.018	10.69 (3.54)	0.416	12.42 (3.84)	0.363
	Saxenda®	19.54 (4.31)		15.16 (5.77)		12.65 (2.38)		11.20 (3.69)		11.45 (3.96)		13.19 (4.30)	
	Mounjaro®	19.90 (3.55)		16.39 (5.25)		12.41 (2.50)		11.19 (3.32)		10.62 (3.99)		13.47 (3.60)	

EF = Enjoyment of food, EOE = Emotional overeating, FF = Food fussiness, H = Hunger, SE = Slowness in eating, SR = Satiety responsiveness, GLP-1RA = Glucagon-like peptide-1 receptor agonists. Data are presented as mean ± standard deviation (SD). Kruskal-Wallis tests were used to compare the differences between more than two groups (e.g., obesity categories, educational levels), while Mann-Whitney U tests were used for comparisons between two independent groups (e.g., male vs. female). Statistical significance was set at $p < 0.05$.

outcomes were observed across several demographic categories, particularly with respect to BMI classification, educational level, and anti-obesity medication use.

Gender Differences: There were no statistically significant differences between males and females in psycho-social impact (25.23 vs. 25.98, $p = 0.357$), diet experience (11.82 vs. 12.33, $p = 0.527$), or total QoL score (58.82 vs. 70.23, $p = 0.553$). However, males reported higher physical impact scores (31.91 vs. 21.76, $p = 0.097$), suggesting greater physical impairment related to obesity; although this did not reach statistical significance, it may reflect clinically meaningful differences warranting further exploration.

Obesity Category: Significant declines in all QoL domains were observed with increasing BMI classification. Normal weight individuals had the highest total QoL score (79.27) and lowest physical impact (35.54), while those in the obese class III group experienced the most severe impairments, including the lowest total QoL (52.30) and physical impact (20.02, $p < 0.001$) scores. Similar patterns were seen for psycho-social impact (22.22, $p < 0.001$) and diet experience (10.05, $p < 0.001$). These

findings demonstrate a clear inverse relationship between BMI and health-related QoL, with progressively worse outcomes in more severe obesity classes.

Marital Status: Single participants reported slightly higher physical impact (33.15 vs. 29.37, p = 0.031) compared to married individuals, indicating greater physical burden among single individuals. However, no significant differences were found in psycho-social impact (27.72 vs. 25.55, p = 0.156), diet experience (12.09 vs. 12.25, p = 0.914), or total QoL (72.96 vs. 67.18, p = 0.054), suggesting that marital status has limited influence on overall QoL beyond physical aspects.

Occupation: Occupational status showed minimal effect on QoL outcomes. No significant differences were found in physical impact (p = 0.341), psycho-social impact (p = 0.550), or total QoL (p = 0.235). However, diet experience differed significantly across occupational groups (p = 0.006), with non-working individuals reporting the highest diet experience scores (12.72), followed by employers (12.22) and students (11.40), suggesting that employment status may modestly influence dietary satisfaction or restriction.

Educational Level: Higher educational attainment was associated with better QoL outcomes. Participants with secondary education scored lower in physical impact (25.33, p = 0.001), psycho-social impact (24.94, p = 0.049), and total QoL (62.47, p = 0.009) compared to university-educated participants. Postgraduate individuals had the highest physical impact score (32.15) but comparable psycho-social and total QoL scores to other high-education groups. These findings suggest that educational background may influence both physical and emotional well-being in relation to obesity.

Anti-obesity medication use: Participants who had never used any anti-obesity medications reported the worst QoL outcomes across all domains. Non-users had a significantly higher physical impact (36.16, p = 0.001), psycho-social impact (31.51, p = 0.001), and total QoL (83.48, p = 0.001) scores compared to users of Orlistat, herbal products, or GLP-1RAs. Users of GLP-1RAs had the best overall QoL (67.93), consistent with their role in improving both metabolic and behavioral outcomes.

The type of GLP-1RA Used: No significant differences were observed in QoL domains between users of Ozempic®, Saxenda®, or Mounjaro®. Physical impact (30.56 vs. 28.93 vs. 30.19, p = 0.646), psycho-social impact (24.97 vs. 27.06 vs. 25.25, p = 0.263), diet experience (11.84 vs. 12.32 vs. 12.88, p = 0.263), and total QoL (67.37 vs. 68.32 vs. 68.33, p = 0.900) did not vary significantly across GLP-1RA types, suggesting similar effects on QoL despite potential differences in appetite regulation and behavioral traits previously observed (Table 6).

Table 6 Difference in quality-of-life measures between each demographic group

Demographic variable		Physical impact		Psycho-social impact		Diet experience		Total QOLOD	
		Mean (SD)	P value	Mean (SD)	P value	Mean (SD)	P value	Mean (SD)	P value
Gender	Male	31.91 (8.11)		25.23 (7.13)		11.82 (3.72)		58.82 (18.09)	
	Female	21.76 (10.34	0.097	25.98 (8.72)	0.357	12.33 (3.98)	0.527	70.23 (23)	0.553
Obesity category	Normal weight	35.54 (7.67)		30.72 (8.12)		13.00 (2.72)		79.27 (15.68)	
	Overweight	34.45 (8.38)	0.001	28.10 (7.74)	0.001	13.09 (3.91)	0.001	75.65 (15.34)	0.001
	Obese class I	32.26 (7.76)		26.00 (8.81)		12.60 (4.32)		70.86 (17.87)	
	Obese class II	26.68 (8.48)		24.31 (8.02)		12.03 (3.35)		63.03 (15.53)	
	Obese class III	20.02 (7.43)		22.22 (7.89		10.05 (3.52)		52.30 (13.58)	
Marital status	Single	33.15 (7.71)	0.031	27.72 (8.74)	0.156	12.09 (3.93)	0.914	72.96 (16.63)	0.054
	Married	29.37 (9.66)		25.55 (8.35)		12.25 (3.93)		67.18 (17.93)	
Occupation	Student	29.21 (11.08)		25.68 (7.38)		11.40 (4.81)		66.31 (18.74)	
	Employer	30.14 (9.32)	0.341	25.98 (8.43)	0.550	12.22 (3.78)	0.006	68.35 (17.38)	0.235
	Not working	29.38 (9.26)		25.45 (9.04)		12.72 (3.80)		67.56 (19.00)	
Educational level	Primary education	29.33 (6.77)		26.83 (8.49)		12.50 (3.39)		68.66 (12.37)	
	Secondary education	25.33 (10.13)	0.001	24.94 (9.06)	0.049	12.19 (4.57)	0.305	62.47 (20.13)	0.009
	University education	30.08 (9.53)		26.27 (8.04)		12.46 (3.72)		68.82 (17.88)	
	Postgraduate education	32.15 (8.47)		25.05 (9.11)		11.59 (4.12)		68.80	
Obesity drug used	No drug	36.16 (8.42)		31.51 (8.25)		31.51 (8.25)		83.48 (17.56)	
	Orlistat	31.55 (10.36)	0.001	27.25 (9.38)	0.001	27.25 (9.38)	0.001	72.50 (20.97)	0.001
	Herbals	31.84 (8.84)		26.54 (8.48)		26.54 (8.48)		71.16 (19.30)	
	GLP-1RAs	29.86 (9.51)		25.83 (8.41)		25.83 (8.41)		67.93 (17.85)	
GLP-1RAs using or not	Ozempic®	30.56 (8.76)		24.97 (8.80)		11.84 (4.00)		67.37 (17.52)	
	Saxenda®	28.93 (9.97)	0.646	27.06 (8.03)	0.263	12.32 (3.83)	0.263	68.32 (18.03)	0.900
	Mounjaro®	30.19 (10.11)		25.25 (8.16)		12.88 (3.94)		68.33 (18.48)	

EF = Enjoyment of food, EOE = Emotional overeating, FF = Food fussiness, H = Hunger, SE = Slowness in eating, SR = Satiety responsiveness, GLP-1RA = Glucagon-like peptide-1 receptor agonists, QOLOD = Quality of Life, Obesity and Dietetics. Data are presented as mean ± standard deviation (SD). Kruskal-Wallis tests were used to compare the differences between more than two groups (e.g., obesity categories, educational levels), and Mann-Whitney U tests were used for comparisons between two independent groups (e.g., male vs. female). Statistical significance was set at p < 0.05.

Correlations Between BMI and Behavioral and Quality of Life Scores

Spearman correlation analysis revealed several statistically significant associations between BMI and appetitive traits, as well as health-related QoL domains. EOE showed a modest but significant positive correlation with BMI (rs = +0.194, p < 0.001), indicating that individuals with higher BMI tended to experience greater EOE behaviors. In contrast, slowness in eating (rs = –0.117, p = 0.003) and satiety responsiveness (rs = –0.112, p = 0.001) were negatively correlated with BMI, suggesting that individuals with higher BMI exhibited reduced sensitivity to satiety cues and slower eating pace.

Enjoyment of food (rs = +0.057, p = 0.066) and food fussiness (rs = –0.018, p = 0.575) did not show statistically significant correlations with BMI. Hunger also demonstrated a weak and non-significant negative association with BMI (rs = –0.056, p = 0.080).

Table 7 Spearman Correlations Between BMI and Behavioral Scores

Subscale	r_s	p-value
Enjoyment of food	+0.057	0.066
Emotional overeating	+0.194	< 0.001
Food fussiness	-0.018	0.575
Hunger	-0.056	0.080
Slowness in eating	-0.117	0.003
Satiety responsiveness	-0.112	0.001
Physical Impact	-0.322	< 0.001
Psycho-social Impact	-0.264	< 0.001
Diet Experience	-0.223	< 0.001

r_s: Spearman correlation coefficient. Spearman's rank correlation (r_s) was used to assess the strength and direction of the relationship between BMI and each behavioral score. A p-value of < 0.05 indicates a statistically significant correlation. Spearman's r_s is used as it is appropriate for non-normally distributed data, providing a measure of association for continuous and ordinal variables.

In terms of QoL, all subscales showed significant inverse correlations with BMI, consistent with a decline in QoL as BMI increased. Physical impact had the strongest negative correlation with BMI (rs = –0.322, p < 0.001), followed by psycho-social impact (rs = –0.264, p < 0.001) and diet experience (rs = –0.223, p < 0.001). These results indicate that individuals with higher BMI experienced greater physical limitations, poorer psycho-social well-being, and more adverse dietary experiences (Table 7).

These findings underscore the complex relationship between BMI, eating behavior traits, and health-related QoL. The observed correlations support the hypothesis that obesity is not only a metabolic condition but also one that is closely linked with altered appetitive behaviors and substantial impairment in multiple domains of daily functioning and well-being.

Behavioral Clustering Analysis

K-means clustering was performed to identify distinct behavioral subgroups based on key appetitive traits. The analysis revealed two primary clusters: cluster 1 ("fuzzy eaters") and cluster 2 ("gourmets"), characterized by distinct patterns in standardized z-scores for eating behaviors (Figure 1). These clusters were further explored in relation to demographic variables and their implications for obesity management.

The final cluster centers (Figure 1) illustrated significant differences between the two clusters: Fuzzy eaters exhibited significantly lower enjoyment of food (Z-score ≈ –0.6) compared to gourmets. Demonstrated higher EOE (Z-score ≈ 0.6) and hunger (Z-score ≈ 0.5). Showed reduced satiety responsiveness (Z-score ≈ –0.6) and slowness in eating (Z-score ≈ –0.3). Food fussiness was slightly elevated but not as pronounced as other traits (Z-score ≈ 0.2).

Gourmets displayed higher enjoyment of food (Z-score ≈ 0.6) compared to Fuzzy eaters. Had lower EOE (Z-score ≈ 0.5) and hunger (Z-score ≈ –0.2). Exhibited greater satiety responsiveness (Z-score ≈ 0.5) and slowness in eating (Z-score ≈ –0.2). Food fussiness was relatively low (Z-score ≈ –0.1).

These findings suggest that fuzzy eaters are more prone to impulsive or "fuzzy" eating behaviors, characterized by heightened EOE, reduced satiety, and increased hunger. In contrast, gourmets individuals exhibit more controlled eating behaviors, marked by higher enjoyment of food, better satiety responsiveness, and slower eating pace.

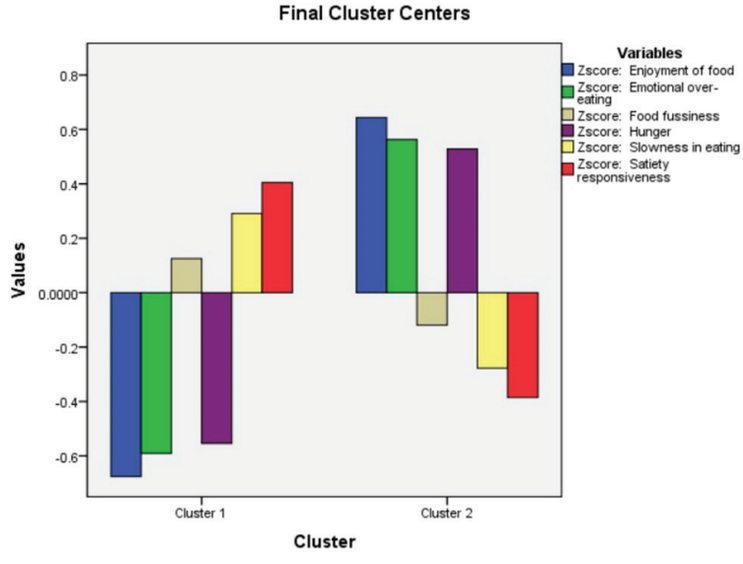


Figure 1 – Final clusters using K means clustering

Behavioral Clustering and Demographic Associations

K-means clustering identified two distinct behavioral subgroups based on appetitive traits: fuzzy eaters and gourmets. The distribution of these clusters across demographic and clinical variables revealed statistically significant differences in gender, BMI classification, educational level, obesity drug use, and type of GLP-1RA used.

Gender Differences: There was a significantly higher proportion of males in fuzzy eaters' cluster compared to gourmets' cluster (86 vs. 63, p = 0.021). In contrast, females were more evenly distributed between the two clusters (399 in fuzzy eaters vs. 446 in gourmets), suggesting that male participants were more likely to exhibit impulsive or less regulated eating behaviors associated with fuzzy eaters' cluster.

Obesity Category: BMI classification showed a strong association with cluster membership. Normal weight individuals were more frequently represented in gourmets compared to overweight and obese groups (fuzzy eaters: 157 vs. gourmets: 82, p = 0.001). Conversely, the proportion of individuals in fuzzy eaters' cluster increased with higher BMI categories, particularly among those classified as obese class I and II, indicating that disordered eating patterns were more prevalent in individuals with greater body weight.

Table 8 Association between clusters and the demographic characteristics

Variable		Cluster 1 (fuzzy eaters)	Cluster 2 (gourmets)	P- value
Gender	Male	86	63	0.021
	Female	399	446	
Obesity category	Normal weight	157	82	0.001
	Overweight	154	186	
	Obese class I	90	133	
	Obese class II	55	67	
	Obese class III	29	41	
Marital status	Single	109	93	0.115
	Married	376	416	
Occupation	Student	63	71	0.83
	Employer	260	301	
	Not working	162	137	
Educational level	Primary education	13	7	0.002
	Secondary education	69	51	
	University education	340	346	
	Postgraduate education	63	105	
Obesity drug used	No drug	330	301	0.006
	Orlistat	17	23	
	Herbals	21	45	
	GLP-1RAs	117	140	
Type of GLP-1RAs used	Ozempic®	39	68	0.042
	Saxenda®	50	49	
	Mounjaro®	28	23	

The Chi-square test was used to assess the association between demographic characteristics (e.g., gender, obesity category, marital status, etc.) and the two clusters (Fuzzy Eaters vs. Gourmets). A p-value of < 0.05 indicates a significant association between the variables.

Educational Level: Educational attainment was also linked to cluster assignment. Participants with only primary education were more likely to belong to fuzzy eaters' cluster compared to those with postgraduate education (13 vs. 7, p = 0.002). Higher levels of education, especially postgraduate training, were associated with gourmets' cluster, suggesting that better-educated individuals tended to display healthier eating behaviors.

Obesity Drug Use: The pattern of anti-obesity medication use differed significantly between clusters. Non-users of anti-obesity drugs were more commonly found in fuzzy eaters' cluster than in gourmets' cluster (330 vs. 301, p = 0.006). Among users, individuals taking herbal products were more frequently assigned to gourmets' cluster, whereas Orlistat users were relatively balanced between clusters. These findings suggest that medication use may be associated with specific eating behavior profiles.

Type of GLP-1RA Used: Among GLP-1RAs users, there was a significant difference in cluster distribution depending on the specific agent used. Ozempic® users were more likely to belong to gourmets' cluster compared to users of Saxenda® or Mounjaro® (Ozempic®: fuzzy eaters' cluster = 39 vs. gourmets' cluster= 68, p = 0.042). This indicates potential differential effects of GLP-1RAs on appetite regulation and eating behavior patterns (Table 8).

These results highlight the heterogeneity in eating behaviors across different demographic and clinical subpopulations. The observed associations support the idea that behavioral clustering can help identify meaningful subgroups for targeted interventions in obesity management.

Multivariable Regression Analysis of Quality of Life

A multiple linear regression model was conducted to identify independent predictors of total QoL score (QoL total), based on demographic characteristics, BMI classification, behavioral clustering, and anti-obesity medication use. The model included ten predictor variables: cluster membership (fuzzy eaters vs. gourmets), gender, BMI category (overweight, obese class I–III), prior use of orlistat or herbal weight loss products, and current use of GLP-1RAs. The regression model demonstrated a statistically significant fit, explaining 34% of the variance in total QoL score ($R^2 = 0.340$, adjusted $R^2 = 0.333$), with an overall model F-statistic of $F(10, 983) = 50.533$, $p < 0.001$, indicating strong explanatory power.

Higher BMI was strongly associated with poorer QoL. Compared to individuals with normal weight, those who were overweight had significantly lower QoL scores ($\beta = –0.158$, $p < 0.001$), with progressively worse outcomes observed in higher obesity classes. Specifically, individuals classified as obese class I had significantly reduced QoL ($\beta = –0.237$, $p < 0.001$), while those in obese class II ($\beta = –0.358$, $p < 0.001$) and class III ($\beta = –0.364$, $p < 0.001$) exhibited the most substantial impairments in overall QoL.

Cluster membership also emerged as a significant predictor. Individuals categorized into fuzzy eaters' cluster reported significantly lower QoL compared to gourmets' cluster ($B = 9.689$, $\beta = 0.252$, $p < 0.001$), highlighting the impact of disordered eating behaviors on health-related QoL. Being male was independently associated with worse QoL ($B = 4.127$, $\beta = 0.077$, $p = 0.005$), suggesting gender-based differences in perceived well-being among participants.

Use of anti-obesity medications showed mixed associations. Current users of GLP-1RAs reported significantly

Table 9 Multiple linear regression coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	84.498	1.286		65.717	0.000		
Cluster Number of Case	9.689	1.189	0.252	8.150	0.000	0.702	1.425
Gender	4.127	1.468	0.077	2.812	0.005	0.903	1.107
BMI group (Overweight)	-6.400	1.375	-0.158	-4.654	0.000	0.582	1.717
BMI group (Obese Class I)	-10.922	1.572	-0.237	-6.949	0.000	0.577	1.734
BMI group (Obese Class II)	-20.934	1.916	-0.358	-10.925	0.000	0.627	1.595
BMI group (Obese Class III)	-27.359	2.332	-0.364	-11.734	0.000	0.697	1.436
Obesity drug (orlistat)	-5.435	2.600	-0.056	-2.091	0.037	0.950	1.053
Obesity drug (herbal)	-3.934	2.102	-0.051	-1.871	0.062	0.905	1.105
GLP-1RAs cluster	-6.740	2.310	-0.113	-2.917	0.004	0.447	2.236
Obesity drug (GLP-1RAs)	-4.757	1.649	-0.108	-2.885	0.004	0.476	2.103

B = Unstandardized Coefficient. Beta = Standardized Coefficient. T = T-statistic. Sig. = P-value (Significance). Std. Error = Standard Error of the Unstandardized Coefficients. Tolerance = Collinearity Tolerance (measures how much a variable is explained by the other predictors). VIF = Variance Inflation Factor (measures the severity of multicollinearity). GLP-1RAs = Glucagon-like peptide-1 receptor agonists. BMI = Body Mass Index. A multiple linear regression model was used to examine the relationships between various predictor variables (e.g., cluster number, gender, BMI categories, obesity drug use, and GLP-1RAs use) and the dependent variable. Unstandardized coefficients (B) represent the change in the dependent variable for a one-unit change in the predictor variable, while standardized coefficients (Beta) show the relative importance of each predictor in the model. Significance: The p-value (Sig.) for each predictor is provided. A p-value of less than 0.05 indicates that the corresponding predictor is statistically significant. For example, Cluster, BMI groups, and GLP-1RAs have significant associations with the dependent variable as indicated by their p-values. Collinearity Statistics: Tolerance values greater than 0.1 and VIF (Variance Inflation Factor) values less than 10 indicate that multicollinearity is not a concern in the model.

better QoL than non-users (B = –4.757, β = –0.108, p = 0.004). Prior use of orlistat was also linked to lower QoL (B = –5.435, β = –0.056, p = 0.037), although this effect was weaker and marginally significant. Use of herbal weight loss products approached statistical significance (B = –3.934, β =–0.051, p = 0.062), indicating a potential negative association between unregulated weight loss methods and QoL.

Collinearity diagnostics revealed acceptable levels of independence between predictor variables, with tolerance values ranging from 0.447 to 0.950 and variance inflation factor (VIF) values below 2.5 across all variables, suggesting that multicollinearity did not compromise the validity of the model.

In summary, this regression analysis identified BMI classification, behavioral cluster, gender, and use of GLP-1RAs as key independent predictors of QoL in individuals across the weight spectrum. These findings reinforce the multifactorial nature of obesity-related QoL and support the integration of both behavioral and pharmacological assessments in clinical approaches to obesity management (Table 9).

Discussion

This study examined the relationship between obesity, eating behaviors, and QoL in Iraqi adults, particularly focusing on the effects of GLP-1RA use. The main findings revealed that GLP-1RA users reported significantly lower hunger levels and better overall QoL compared to non-users, although EOE remained elevated in both groups. Behavioral clustering identified two distinct groups: "fuzzy eaters," who exhibited high EOE and low satiety responsiveness, and "gourmets," who demonstrated better satiety and slower eating. Despite the improvements in appetite regulation observed with GLP-1RA use, EOE was not fully addressed, indicating that pharmacological treatments alone may not completely address the psychological aspects of eating behavior.

These results are consistent with the findings of Charlotte C. van Ruiten et al. [11], who also observed that EOE persisted in patients using GLP-1RAs, despite improvements in physical health outcomes. This suggests that while GLP-1RA treatment can reduce hunger and improve appetite control, EOE remains a significant barrier to achieving optimal outcomes. Similarly, Blundell et al. [14] and Davies et al. [15] reported reductions in EOE scores with GLP-1RA treatment, although these effects were not uniform across all participants, reinforcing the notion that EOE is multifactorial and requires more individualized treatment approaches. This supports our hypothesis that EOE may modulate the efficacy of GLP-1RA treatment, and additional interventions—especially psychological ones—may be necessary for better management of EOE.

In contrast to these findings, Suzuki et al. [16] demonstrated that tirzepatide, a GLP-1RA, significantly improved EOE and glycemic control in their clinical trial participants. This discrepancy may arise due to differences in study design, the

duration of treatment, and the specific GLP-1RA used. Suzuki et al.'s study involved a longer treatment period, which may have allowed for more profound changes in both physical and psychological aspects of eating behavior. The longer treatment duration in their study may account for the more significant impact on EOE compared to our cohort, where the treatment duration was shorter. Moreover, the cultural context of eating behaviors in Iraq, where EOE is often tied to social gatherings and cultural norms around food, may influence responses to treatment. The cultural differences in dietary habits and social food-related behaviors could partly explain why some studies report more significant effects of GLP-1RAs on EOE than others.

An important contribution of this study is the identification of distinct behavioral profiles—"fuzzy eaters" and "gourmets." These two groups exhibited different eating behaviors that were not solely determined by physiological hunger and satiety cues but also by emotional factors. "Fuzzy eaters," with higher EOE tendencies, may be more likely to use food as a coping mechanism for emotional distress. In contrast, "gourmets," who exhibited slower eating and better satiety responsiveness, seem to rely more on internal physiological cues and have fewer issues with EOE. This finding aligns with van Strien et al. [17], who found similar patterns of eating behavior in obese populations, where EOE and external eating were common. The "fuzzy eater" and "gourmet" profiles identified in our study offer valuable insight into how different individuals with obesity might respond to treatment. It suggests that tailored interventions are necessary to address specific behavioral profiles, which could lead to better treatment outcomes for patients.

The role of GLP-1RA treatment in addressing eating behaviors is further complicated by the psychological and emotional factors influencing overeating. As indicated in earlier research, including studies by Shaw et al. [18] and van Strien et al. [17], EOE is a significant challenge in obesity treatment. While GLP-1RAs are effective in controlling appetite and promoting weight loss, they do not fully address the emotional aspects of eating. The persistence of EOE in our study's participants, despite the use of GLP-1RAs, reinforces the idea that additional interventions—such as cognitive-behavioral therapy (CBT) or other psychological approaches—may be necessary to address the emotional drivers of overeating.

Obesity is not just about weight or appetite—it involves metabolic and inflammatory changes that begin early in life. A recent study in Kazakh children found that sleep disturbances are linked to insulin resistance and higher levels of inflammation, even at a young age [19]. While our study focused on adults and did not measure biomarkers, this supports the idea that obesity affects multiple body systems. The persistence of EOE in our GLP-1RA users, despite treatment, may reflect similar underlying dysregulation that goes beyond appetite control. This finding highlights the multifactorial nature of obesity, where hormonal, psychological, and metabolic factors are interconnected, making comprehensive treatment approaches crucial.

The cultural context also plays a pivotal role in shaping eating behaviors. In Iraq, food is central to social interactions, and traditional meals are often associated with strong emotional and social ties. This cultural backdrop may contribute to higher levels of EOE, as food-related behaviors are often intertwined with social experiences. These findings are consistent with Wardle et al. [20] and Raspopow et al. [21], who emphasized the influence of cultural beliefs and social settings on eating habits. Culturally tailored interventions that address both physiological and psychological components of eating may be more effective in helping individuals manage EOE and improve weight management outcomes.

Our findings are consistent with previous studies on behavioral patterns in people with obesity. Research by van Strien et al. [17] and Herman et al. [22] has shown that emotional eaters and external eaters tend to have more difficulties with overeating and managing their weight. This suggests that treatment plans should be customized to address specific eating habits. The identification of 'fuzzy eaters' and 'gourmets' in our study further highlights the importance of personalizing treatment approaches. Doing so could enhance the effectiveness of GLP-1RAs and potentially improve outcomes for those living with obesity.

Despite the valuable insights provided by this study, several limitations should be noted. First, the cross-sectional design of the study prevents causal conclusions about the effects of GLP-1RA use on EOE and QoL. Longitudinal studies are needed to examine the long-term effects of GLP-1RAs on EOE and to determine whether the improvements in appetite regulation observed in this study are sustained over time. Additionally, the reliance on self-reported measures, such as the AEBQ and QOLOD, introduces the potential for response bias, especially regarding EOE, which is a subjective and sensitive measure. Although efforts were made to anonymize survey responses and reduce social desirability bias, future studies could benefit from using more objective measures, such as food diaries or direct observation, to obtain more accurate data.

Another limitation of the study is the lack of diversity in the sample. All participants were recruited from healthcare centers and dietetics clinics in Iraq, which may not fully represent the general population. Including a more diverse sample in future studies would increase the generalizability of the findings. Additionally, while we controlled for several potential confounders in the analysis, other unmeasured factors—such as stress, mental health status, and social support—may influence eating behaviors and QoL. These factors were not assessed in this study but could provide valuable insights into the psychological components of eating behavior.

Finally, future research should explore the potential benefits of combining GLP-1RA treatment with psychological therapies, such as CBT, to address the emotional aspects of eating. CBT has been shown to be effective in treating EOE and could complement the physiological benefits of GLP-1RAs, leading to more sustainable weight management outcomes. Long-term, multi-component interventions that include pharmacological treatments and psychological support may offer the most comprehensive approach to managing obesity and improving QoL in individuals with EOE tendencies.

Conclusions

This study demonstrates that while GLP-1RAs show promise in improving appetite control and QoL in individuals with obesity, they do not fully address EOE, which remains a significant challenge. The identification of distinct behavioral profiles—such as "fuzzy eaters" and "gourmets"—suggests that personalized treatment strategies are needed to target specific behaviors and improve outcomes. Future research should explore the long-term effects of GLP-1RA use and investigate integrated approaches combining pharmacological treatments

with behavioral therapies to better address the psychological components of obesity.

Author Contributions: Conceptualization, M.I.A.; methodology, A.S.A. and M. I.A.; validation, M.I.A.; formal analysis, A.S.A. and M. I.A.; investigation, A.S.A. and M. I.A.; resources, A.S.A. and M. I.A.; data curation, A.S.A. and M. I.A.; writing – original draft preparation, A.S.A. and M. I.A.; writing – review and editing, M. I.A.; visualization, A.S.A. and M. I.A.; supervision, M. I. A.; project administration, M. I.A.; funding acquisition, not applicable. All authors have read and agreed to the published version of the manuscript

Disclosures: The authors have no conflicts of interest.

Acknowledgments: The authors acknowledge the dietitians helped in recruiting the participants. Furthermore, the authors thanks all the survey participants for their valuable contribution.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. Bray GA, Frühbeck G, Ryan DH, Wilding JPH. Obesity: a chronic relapsing progressive disease process. A position statement of the World Obesity Federation. Obes Rev. 2017; 18 (7): 715–723. https://doi.org/10.1111/obr.12551.
2. World Health Organization. Obesity and overweight. World Health Organization. 2024. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight .
3. Global Nutrition Report. Country Nutrition Profiles. *Global Nutrition Report.* 2025. https://globalnutritionreport.org/resources/nutrition-profiles/asia/western-asia/iraq/.
4. Al-Shammaa ZM, Aladul MI, Essa NS. Assessment of appetitive traits among Iraqi population. *Clin Epidemiol Glob Health.* 2025; 33: 102014. https://doi.org/10.1016/j.cegh.2025.102014.
5. Webber J. The comorbidities of obesity. *Pract Diabetes Int.* 2001; 18 (8): 293–296. https://doi.org/10.1002/pdi.274.
6. Katz DA, McHorney CA, Atkinson RL. Impact of obesity on health-related quality of life in patients with chronic illness. *J Gen Intern Med*. 2000; 15 (11): 789–796. https://doi.org/10.1046/j.1525-1497.2000.90906.x.
7. van Strien T, Frijters JER, Bergers GPA, Defares PB. The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *Int J Eat Disord.* 1986; 5 (2): 295–315. https://doi.org/10.1002/eat.20424.
8. Boutelle KN, Manzano MA, Eichen DM. Appetitive traits as targets for weight loss: The role of food cue responsiveness and satiety responsiveness. *Physiol Behav*. 2020; 224: 113018. https://doi.org/10.1016/j.physbeh.2020.113018.
9. Wilding JPH, Batterham RL, Calanna S, Davies M, Van Gaal LF, Lingvay I, McGowan BM, Rosenstock J, Tran MBD, Wadden TA, Wharton S, Yokote K, Zeuthen N, Kushner RF. Once-Weekly Semaglutide in Adults with Overweight or Obesity. *N Engl J Med.* 2021; 384 (11): 989–1002. https://doi.org/10.1016/j.physbeh.2020.113018.
10. Furlan GB, de Souza RJ, Mancini MC, Halpern B. Use of GLP-1 analogs in the treatment of obesity: an integrative and systematic review. *Int J Nutrology.* 2023; 16 (1). https://doi.org/10.54448/ijn23105.
11. Charlotte C. van Ruiten, Jennifer S. ten Kulve, Liselotte van Bloemendaal, Max Nieuwdorp, Dick J. Veltman, Richard G. IJzermanvan. Eating behavior modulates the sensitivity to the central effects of GLP-1 receptor agonist treatment: a secondary analysis of a randomized trial. *Psychoneuroendocrinology*. 2022; 137: 105667. https://doi.org/10.1016/j.psyneuen.2022.105667.
12. Alruwaitaa MA, Alenezi A, Almutairi A, Almutairi A, Alotaibi A, Alharbi A, Aljaser A, Aljaser F, Alzahrani A, Alzahrani F, Alzahrani M, Alzahrani N, Alzahrani R, Alzahrani S, Alzaid A. The Arabic Version of the Adult Eating Behavior Questionnaire among Saudi Population: Translation and Validation. *Nutrients.* 2022; 14 (21): 4705. https://doi.org/10.3390/nu14214705.
13. Ziegler O, Scheen A, Nouwen A, Rigalleau F, Pichelin M, Poulain JP, Hanachi M, Fournier G, Roche F, Drouin P, Lantieri O, Taillard J, Philip P. Development and validation of a French obesity-specific quality of life questionnaire: Quality of Life, Obesity and Dietetics (QOLOD) rating scale. *Diabetes Metab.* 2005; 31(3 Pt 1): 273–283. https://doi.org/10.1016/s1262-3636(07)70194-5.
14. Blundell JE, Finlayson G, Axelsen M, Flint A, Gibbons C, Kvist T, Lawton C, Martins C, Moller EC, Nordgren TM, Pedersen SD, Radionov A, Raben A, Sorensen LB, Westerterp-Plantenga MS, Winkler J, Zhang Y, Chan PS. Effects of once-weekly semaglutide on appetite, energy intake, control of eating, food preference and body weight in subjects with obesity. *Diabetes Obes Metab.* 2017; 19 (9): 1242–1251. https://doi.org/10.1111/dom.12932.
15. Davies MJ, Bergenstal R, Bode B, Kushner RF, Lewin A, Skjøth TV, Harney K, Koenen C, Bain SC. Efficacy of Liraglutide for Weight Loss Among Patients With Type 2 Diabetes: The SCALE Diabetes Randomized Clinical Trial. *JAMA*. 2015; 314 (7): 687–699. https://doi.org/10.1001/jama.2015.9676.
16. Suzuki T, Ito C, Kato Y, Kusunoki Y, Kawai K, Murai M, Seino Y, Inagaki N. Tirzepatide ameliorates eating behaviors regardless of prior exposure to glucagon-like peptide receptor agonists in Japanese patients with type 2 diabetes mellitus. *J Diabetes Complications.* 2024; 38 (7): 108779. https://doi.org/10.1016/j.jdiacomp.2024.108779.
17. van Strien T, van Beers I, van Kruistum C, Hakko E, Herman CP, Anschutz D. Emotional eating and food intake after sadness and joy. Appetite. 2013; 66: 20–25. https://doi.org/10.1016/j.appet.2013.02.016.
18. Shaw K, O'Rourke P, Del Mar C, Kenardy J. Psychological interventions for overweight or obesity. *Cochrane Database Syst Rev.* 2005 Apr 18;(2):CD003818. https://doi.org/10.1002/14651858.CD003818.pub2.

19. Konuksever D, Kaya AK, Çelik N, Öztürk A, Karakaş Ç, Karakuş E, Yılmaz S, Öztürk Y. The association between insulin resistance, sleep disorders, and inflammation in obese children. *J Clin Med Kaz.* 2023;20(3):82–87. https://doi.org/10.23950/jcmk/13327.
20. Wardle J, Haase AM, Steptoe A, Nillapun M, Jonwutiwes K, Bellisle F. Gender differences in food choice: the contribution of health beliefs and dieting. *Ann Behav Med.* 2004; 27 (2): 107–116. https://doi.org/10.1207/s15324796abm2702_5.
21. Raspopow K, Matheson K, Abizaid A, Anisman H. Unsupportive social interactions influence emotional eating behaviors. The role of coping styles as mediators. *Appetite*. 2013 Mar;62:143-9. https://doi.org/10.1016/j.appet.2012.11.031.
22. Herman CP, Polivy J. External cues in the control of food intake in humans: the sensory-normative distinction. *Physiol Behav.* 2008 Aug 6;94(5):722-8. https://doi.org/10.1016/j.physbeh.2008.04.014.

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17414

Rapid Hepatic Function Improvement with Culturally Adapted Low-Carbohydrate Paleolithic Diet in South Asian Adults with Metabolic Dysfunction: A 12-Week Prospective Study with Predictive Modeling

Hariharan Venkatasubramanian[1], Vijayasamundeeswari Chinnathambipalayam Kandaswamy[2]

[1]Department of Biochemistry, Vinayaka Mission's Kirupananda Variyar Medical College, Vinayaka Mission's Research Foundation (Deemed to be University), Tamil Nadu, India

Received: 2025-08-22.
Accepted: 2025-11-11.

J Clin Med Kaz 2025; 22(6) 73–79

Corresponding author:
Hariharan Venkatasubramanian.
E-mail: drhariharanv@gmail.com
ORCID: 0000-0002-7937-9070.

ABSTRACT

Background: Non-alcoholic fatty liver disease (NAFLD) affects 25-30% of South Asian adults, with limited predictive tools for dietary intervention success. Identifying baseline predictors of hepatic response could enable personalized treatment approaches.

Objective: To evaluate hepatic function benefits of culturally adapted low-carbohydrate Paleolithic diet in South Asian adults with metabolic dysfunction and develop predictive models for treatment response.

Methods: 103 overweight/obese South Asian adults followed a culturally adapted Paleolithic diet for 12 weeks. Primary outcomes were hepatic function improvements (ALT, ALP, hepatic steatosis index). Multiple regression analysis identified predictors of hepatic response. ROC analysis determined optimal baseline cutoffs for treatment success, defined as ≥20% ALT reduction.

Results: Significant hepatic improvements occurred: ALT decreased 17.1% ($p<0.05$), ALP decreased 7.4% ($p<0.01$), hepatic steatosis index improved 10.7% ($p<0.01$). Multiple regression revealed baseline ALT ($\beta=0.58$, $p<0.001$), insulin resistance ($\beta=0.34$, $p<0.01$), and triglycerides ($\beta=0.28$, $p<0.05$) as independent predictors of hepatic response ($R^2=0.64$). ROC analysis showed baseline ALT >28.5 U/L had 81.3% sensitivity and 74.6% specificity for predicting treatment success (AUC=0.82, $p<0.001$). Concurrent benefits included 9.4kg weight loss and 41% insulin resistance improvement.

Conclusion: Culturally adapted Paleolithic diet produces rapid hepatic function improvements in South Asian adults, with baseline ALT >28.5 U/L predicting optimal treatment response. These findings enable personalized intervention strategies for NAFLD management.

Keywords: Non-alcoholic fatty liver disease, hepatic function, Paleolithic diet, South Asian population, predictive modeling, ROC analysis, personalized nutrition

Introduction

Non-alcoholic fatty liver disease (NAFLD) has emerged as the most common chronic liver condition globally, affecting approximately 25% of the worldwide population and representing a significant public health burden [1]. In South Asian populations, NAFLD prevalence reaches 25-30% of adults, with disease manifestation occurring at lower BMI thresholds compared to other ethnic groups [2]. This "metabolically obese, normal weight" phenotype characteristic of South Asians results in earlier and more severe hepatic dysfunction, often progressing to non-alcoholic steatohepatitis (NASH) and advanced fibrosis [3].

The pathophysiology of NAFLD in South Asian populations involves complex interactions between genetic predisposition, insulin resistance, visceral adiposity, and dietary factors [4]. Traditional South Asian diets, characterized by high carbohydrate intake (60-70% of total calories) and frequent meal patterns, may contribute to hepatic lipogenesis and steatosis development [5]. Current treatment approaches primarily focus on weight loss and metabolic improvement, with limited evidence for population-specific dietary interventions targeting hepatic function.

Scientific Rationale and Knowledge Gap

Despite NAFLD's high prevalence in South Asian populations, several critical knowledge gaps exist:

1. Predictive Modeling Deficit: No validated predictive models exist for identifying South Asian individuals most likely to benefit from dietary interventions, limiting personalized treatment approaches [6].

2. Population-Specific Response Patterns: Limited data exists on baseline characteristics predicting hepatic function improvements in culturally adapted dietary interventions [7].

3. Rapid Response Assessment: Absence of validated biomarkers for early treatment response evaluation within 12 weeks of dietary intervention [8].

4. Cultural Adaptation Evidence: Lack of evidence for hepatic benefits of culturally adapted low-carbohydrate interventions in South Asian populations [9].

Low-carbohydrate diets (LCDs) have demonstrated promising results for NAFLD management through multiple mechanisms: reduced hepatic de novo lipogenesis, enhanced fatty acid oxidation, decreased insulin resistance, and improved hepatic insulin sensitivity [10]. Recent meta-analyses have shown that LCDs can reduce hepatic fat content by 20-30% within 12 weeks, with concurrent improvements in liver enzyme levels [11]. However, individual response variability remains poorly understood, with some patients achieving substantial improvements while others show minimal response [12].

Low-carbohydrate diets reduce hepatic lipogenesis and enhance fatty acid oxidation through several interconnected biochemical mechanisms. Carbohydrate restriction decreases insulin secretion and hepatic insulin signaling, which downregulates the transcription factors sterol regulatory element-binding protein-1c (SREBP-1c) and carbohydrate-responsive element-binding protein (ChREBP). These transcription factors normally activate genes encoding lipogenic enzymes including fatty acid synthase (FAS), acetyl-CoA carboxylase (ACC), and stearoyl-CoA desaturase-1 (SCD1). By suppressing these pathways, carbohydrate restriction reduces de novo lipogenesis—the conversion of excess glucose and fructose into hepatic triglycerides. Simultaneously, reduced insulin levels and lower hepatic glucose availability activate AMP-activated protein kinase (AMPK), which promotes fatty acid oxidation by phosphorylating and inhibiting ACC, thereby reducing malonyl-CoA production. Lower malonyl-CoA levels relieve inhibition of carnitine palmitoyltransferase-1 (CPT-1), the rate-limiting enzyme for mitochondrial fatty acid uptake and β-oxidation. This metabolic shift redirects hepatic metabolism from lipid storage to lipid utilization, resulting in reduced hepatic steatosis and improved liver enzyme profiles. Studies demonstrate that low-carbohydrate diets (30% energy from carbohydrates) produce significantly higher total and plasma fatty acid oxidation rates (4.8 and 4.6 μmol/kg/min) compared to high-carbohydrate diets (2.4 and 2.1 μmol/kg/min respectively), while simultaneously suppressing hepatic de novo lipogenesis by up to 75%.

Machine learning approaches and predictive modeling have emerged as valuable tools for personalizing dietary interventions. ROC analysis and multiple regression modeling can identify baseline predictors of treatment success, enabling clinicians to optimize patient selection and intervention strategies [13].

This study addresses these knowledge gaps by evaluating hepatic function benefits of a culturally adapted Paleolithic diet in South Asian adults while developing predictive models for treatment response using advanced statistical approaches including multiple regression and ROC analysis.

Methods

This prospective, interventional study was conducted between March 2022 and February 2023. The study protocol was approved by the Institutional Human Ethics Committee (Reference: IHEC/230/Biochemistry/06/2022) and conducted in accordance with the Declaration of Helsinki.

Participants and Selection Criteria South Asian adults aged 18-65 years with BMI ≥25 kg/m² (South Asian cutoff for obesity) were recruited through consecutive sampling from the institutional diet clinic. Participants with evidence of metabolic dysfunction (defined as presence of at least one: elevated fasting glucose, insulin resistance, dyslipidemia, or elevated liver enzymes) were prioritized for inclusion.

Inclusion criteria: (1) BMI ≥25 kg/m², (2) evidence of metabolic dysfunction, (3) elevated ALT and/or ALP levels, (4) willingness to adhere to dietary intervention for 12 weeks, (5) no current dietary interventions, and (6) provision of informed consent.

Exclusion criteria: (1) established liver disease (hepatitis B/C, autoimmune hepatitis, Wilson's disease), (2) alcohol consumption >20g/day, (3) hepatotoxic medication use, (4) insulin-dependent diabetes mellitus or uncontrolled diabetes requiring pharmacotherapy (HbA1c >9% with ongoing medication), (5) established cardiovascular disease, (6) uncontrolled hypertension, (7) chronic kidney disease, (8) pregnancy/lactation, and (9) acute illness. Participants with prediabetes (HbA1c 5.7-6.4%) and diet-controlled early hyperglycemia were included, as dietary intervention represents first-line management for these conditions. Those requiring insulin or multiple oral hypoglycemic agents were excluded to avoid confounding effects of glucose-lowering medications on hepatic function parameters.

Sample Size and Statistical Power Based on previous studies showing mean ALT reduction of 12±8 U/L with dietary interventions, targeting 80% power at 5% significance level, the calculated sample size was 85 participants. For regression analysis with 6 predictors, minimum sample size of 90 was

required (15 subjects per predictor). Accounting for dropout, 103 participants were enrolled.

Intervention Protocol The culturally adapted Paleolithic diet was designed by registered dietitians with expertise in South Asian cuisine and hepatic nutrition. The intervention emphasized foods traditionally consumed in South Asian populations while restricting hepatogenic carbohydrates.

Macronutrient Distribution: 25% carbohydrates (40-50g daily), 15% protein, 60% healthy fats, with emphasis on hepatoprotective nutrients including omega-3 fatty acids, antioxidants, and fiber.

Permitted foods: Lean meats, fish (especially omega-3 rich varieties), eggs, dairy products (paneer, cheese, ghee), nuts, seeds, non-starchy vegetables, limited low-glycemic fruits (guava, coconut).

Restricted foods: Grains (wheat, rice, millets), legumes, refined sugars, processed foods, high-fructose fruits, and hepatotoxic additives.

Individual caloric requirements were calculated using the Mifflin-St Jeor equation with activity factors. Iso-caloric meal plans were provided to ensure adequate nutrition while targeting hepatic function improvement.

Biochemical Analysis and Equipment

All biochemical parameters were assessed at baseline and 12-week endpoint following an overnight fast (10-12 hours). Blood samples (10 mL venous blood) were collected in EDTA and plain vacutainer tubes under aseptic conditions by trained phlebotomists.

Hepatic Function Parameters: Serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and total bilirubin were measured using kinetic enzymatic methods on an automated clinical chemistry analyzer (Beckman Coulter AU680 Chemistry Analyzer, Brea, CA, USA). ALT and AST were measured by IFCC (International Federation of Clinical Chemistry) kinetic UV method without pyridoxal phosphate activation at 37°C. ALP was measured by pNPP (p-nitrophenyl phosphate) kinetic method at 37°C. Total bilirubin was measured by diazo method (Jendrassik-Grof modified method). Quality control was performed using manufacturer-supplied control sera at two levels (normal and pathological ranges) with each batch. Inter-assay coefficient of variation was <3.5% for ALT, <3.2% for AST, <2.8% for ALP, and <4.1% for bilirubin. Intra-assay coefficient of variation was <2.0% for all hepatic parameters.

Glycemic Parameters: HbA1c was measured by high-performance liquid chromatography (HPLC) using a Bio-Rad D-10 Hemoglobin Testing System (Bio-Rad Laboratories, Hercules, CA, USA) with National Glycohemoglobin Standardization Program (NGSP) certification and calibration traceable to the Diabetes Control and Complications Trial (DCCT) reference method. Measurement range: 3.1-18.5% with CV <2%. Fasting plasma glucose was measured by glucose oxidase-peroxidase (GOD-POD) enzymatic colorimetric method with linearity range 10-600 mg/dL. Serum insulin was quantified using chemiluminescent immunoassay (CLIA) on ADVIA Centaur XP Immunoassay System (Siemens Healthcare Diagnostics, Tarrytown, NY, USA) with analytical sensitivity of 0.5 μIU/mL and functional sensitivity of 1.0 μIU/mL. HOMA-IR (Homeostatic Model Assessment for Insulin Resistance) was calculated using the formula: [fasting insulin (μU/mL) × fasting glucose (mg/dL)] / 405.**

Lipid Profile: Total cholesterol, triglycerides, HDL-cholesterol, and LDL-cholesterol were measured using enzymatic colorimetric methods on the Beckman Coulter AU680 automated analyzer. Total cholesterol was measured by CHOD-PAP (cholesterol oxidase-phenol aminophenazone peroxidase) method, triglycerides by GPO-PAP (glycerol-3-phosphate oxidase-phenol aminophenazone peroxidase) method, and HDL-cholesterol by direct enzymatic colorimetric method with polyethylene glycol-modified enzymes. LDL-cholesterol was calculated using the Friedewald equation for samples with triglycerides <400 mg/dL. All lipid assays were calibrated against CDC-certified reference materials. Inter-assay CVs were: total cholesterol <2.5%, triglycerides <3.0%, HDL <2.8%, LDL <3.2%. Internal quality control was maintained using commercial control sera (Bio-Rad Laboratories) at two levels (normal and abnormal ranges) with each analytical run. External quality assurance was maintained through participation in the External Quality Assessment Scheme (EQAS) conducted by the Christian Medical College, Vellore, India.

Anthropometric Measurements: Body weight was measured using a calibrated digital electronic weighing scale (Omron HBF-212, Omron Healthcare, Kyoto, Japan; accuracy ±0.1 kg, capacity 150 kg) with participants in light clothing without footwear, after voiding, in the morning hours. Height was measured using a wall-mounted stadiometer (Seca 206, Seca GmbH, Hamburg, Germany; accuracy ±0.1 cm, range 70-205 cm) with the participant standing erect, heels together, and head in the Frankfort horizontal plane. Body Mass Index (BMI) was calculated as weight (kg) divided by height squared (m^2).

All measurements were performed by trained laboratory technicians who were blinded to participant grouping and treatment response status. All equipment was calibrated daily according to manufacturer specifications. Hepatic Steatosis Index (HSI) was calculated using the validated formula: HSI = 8 × (ALT/AST ratio) + BMI (+2 if female, +2 if diabetes mellitus present).

Table 1 Culturally Adapted Paleolithic Diet Plan

Meal	Group	Food Items	Hepatic Benefits
Breakfast	All groups	50-80 almonds OR 50-100g peanuts OR 100g coconut + 30g butter in 200ml milk/tea	Vitamin E, healthy fats, reduced fructose
Lunch	Non-Veg/ Ovo-veg	4 whole eggs + 200g vegetables + 30g cheese	High-quality protein, choline, antioxidants
	Vegetarian	300g vegetables + 50g cheese + 50g coconut	Fiber, antioxidants, MCTs
Snacks	All groups	Curd 100ml/guava 100g/ milk 100ml/pistachios 50g, vegetable soups, lemon water	Probiotics, vitamin C, electrolytes
Dinner	Non-Vegetarian	150-300g lean meat/fish (especially fatty fish)	Omega-3, protein, reduced carbs
	Vegetarian	100-200g paneer/ mushrooms + 30g cheese	Protein, reduced hepatic glucose load

Outcome Measures

Primary Outcomes (Hepatic Function):

1. Alanine aminotransferase (ALT) change
2. Alkaline phosphatase (ALP) change
3. Aspartate aminotransferase (AST) change
4. Total bilirubin change

5. Hepatic steatosis index (HSI) = 8×(ALT/AST) + BMI (+2 if female, +2 if diabetes)

Hepatic treatment success defined as ≥20% ALT reduction, based on clinical significance thresholds from hepatology literature [14].

Response Classification and Non-Responder Definition

Participants were classified as 'Responders' or 'Non-responders' based on hepatic treatment success criteria defined a priori. Responders were defined as participants who: (1) completed the full 12-week dietary intervention with documented adherence verified through bi-weekly dietary counseling sessions, food diary reviews, and 24-hour dietary recall interviews conducted by registered dietitians; AND (2) achieved ≥20% reduction in serum ALT levels from baseline to week 12. The 20% threshold was selected based on clinically significant ALT reductions reported in hepatology literature for NAFLD dietary interventions.

Non-responders were defined as participants who: (1) completed the full 12-week intervention with documented dietary adherence meeting the same criteria as responders (≥80% adherence to meal plans based on food diary analysis and bi-weekly counseling records), BUT (2) failed to achieve the ≥20% ALT reduction threshold. Importantly, all 103 participants included in the final analysis demonstrated satisfactory dietary adherence for the complete 12-week period. Seven participants who discontinued the intervention prematurely (before week 8) due to personal reasons (n=4), inability to maintain dietary adherence (n=2), or loss to follow-up (n=1) were excluded from both responder and non-responder classifications and are not included in this analysis.

Therefore, 'Non-responders' in this study specifically refers to participants who maintained documented dietary compliance and completed the full intervention protocol but did not demonstrate the predetermined biochemical response threshold (≥20% ALT reduction), rather than participants with poor adherence or intervention discontinuation. This distinction is critical for the validity of predictive modeling, as it identifies biological, metabolic, and baseline characteristic factors associated with differential treatment response independent of adherence behaviors. This approach enables the predictive model to focus on physiological predictors of response rather than behavioral compliance factors.

Secondary Outcomes:

1. Weight loss and BMI reduction
2. Glycemic control (HbA1c, insulin resistance)
3. Lipid profile improvements

Statistical Analysis

All statistical analyses were performed using SPSS version 28.0 (IBM Corp., Armonk, NY, USA) and R software version 4.3.0 (R Foundation for Statistical Computing, Vienna, Austria). The significance level was set at $\alpha = 0.05$ for all tests, with Bonferroni correction applied for multiple comparisons where appropriate. Continuous variables expressed as means±standard deviations. Paired t-tests compared pre-post intervention values. Stepwise multiple regression identified independent predictors of hepatic response (% ALT reduction). Candidate predictors included baseline: age, BMI, ALT, AST, ALP, HbA1c, insulin, HOMA-IR, triglycerides, HDL, and total cholesterol. Receiver operating characteristic curves determined optimal baseline cutoffs for predicting treatment success. Area under curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated.

Results

Baseline Characteristics 103 participants (70 males, 33 females) completed the 12-week intervention. Mean age was 42.3±8.7 years with baseline BMI 29.6±2.8 kg/m². Treatment success (≥20% ALT reduction) was achieved by 48 participants (46.6%).

Table 2 Baseline Characteristics and Treatment Response Groups

Characteristic	Overall (n=103)	Responders (n=48)	Non-responders (n=55)	p-value
Age (years)	42.3±8.7	43.2±8.1	41.5±9.2	0.31
BMI (kg/m^2)	29.6±2.8	30.1±2.9	29.2±2.6	0.09
Baseline ALT (U/L)	28.1±18.0	36.8±20.4	20.5±11.2	<0.001
Baseline ALP (U/L)	75.2±21.0	80.3±22.7	70.8±18.5	0.02
HOMA-IR	1.7±1.1	2.0±1.2	1.4±0.9	0.004
Triglycerides (mg/dL)	140±77	161±85	122±64	0.008
HDL (mg/dL)	42.6±10.5	41.2±9.8	43.8±11.0	0.19
HbA1c (%)	6.6±4.4	7.2±4.8	6.1±3.9	0.18
Hepatic Steatosis Index	38.2±6.4	40.1±6.8	36.6±5.7	0.004

Significant improvements were observed across hepatic function parameters. ALT decreased from 28.1±18 to 23.3±13 U/L (17.1% reduction, $p<0.05$). ALP showed substantial reduction from 75.2±21 to 69.6±19 U/L (7.4% reduction, $p<0.01$). The hepatic steatosis index improved significantly from 38.2±6.4 to 34.1±7.2 (10.7% improvement, $p<0.01$).

Responders achieved 38.2±14.6% ALT reduction compared to 1.8±12.4% change in non-responders ($p<0.001$). Responders also showed greater ALP reduction (12.8±8.9% vs 2.7±6.8%, $p<0.001$) and hepatic steatosis index improvement (15.4±7.2% vs 6.8±5.9%, $p<0.001$).

Table 3 Hepatic Function and Metabolic Parameters: Overall and by Response Groups

Parameter	Overall Change	Responders (n=48)	Non-responders (n=55)	p-value
PRIMARY HEPATIC OUTCOMES				
ALT (U/L)	-4.8±15.2	-14.1±7.9	+2.9±8.6	<0.001
ALT % change	-17.1±28.4	-38.2±14.6	+1.8±12.4	<0.001
ALP (U/L)	-5.6±16.4	-10.3±11.2	-1.9±7.8	<0.001
ALP % change	-7.4±21.8	-12.8±8.9	-2.7±6.8	<0.001
HIS	-4.1±6.8	-6.2±4.1	-2.3±3.6	<0.001
SECONDARY OUTCOMES				
Weight loss (kg)	-9.4±4.2	-10.8±3.9	-8.2±4.2	<0.001
BMI reduction	-3.5±1.8	-4.0±1.7	-3.1±1.7	0.006
HbA1c change (%)	-1.3±1.9	-1.7±2.2	-1.0±1.5	0.04
HOMA-IR change	-0.7±1.1	-1.0±1.2	-0.4±0.8	0.003
Triglycerides change (mg/dL)	-35±52	-48±46	-24±37	0.002

Multiple Regression Analysis

Table 4 Multiple Regression Analysis: Predictors of Hepatic Response

Predictor	β Coefficient	Standard Error	t-value	p-value	95% CI
Baseline ALT	0.58	0.08	7.25	<0.001	0.42-0.74
Baseline HOMA-IR	0.34	0.09	3.78	<0.001	0.16-0.52
Baseline Triglycerides	0.28	0.11	2.55	0.012	0.06-0.50
Age	-0.23	0.10	-2.30	0.024	-0.43 to -0.03
BMI	0.19	0.09	2.11	0.038	0.01-0.37

Model Equation: *% ALT Reduction = -18.2 + (0.72 × Baseline ALT) + (6.8 × Baseline HOMA-IR) + (0.08 × Baseline Triglycerides) - (0.28 × Age) + (0.95 × BMI)*
Model Validation: *Cross-validation showed R^2 = 0.61, indicating good model stability.*

ROC Analysis for Treatment Success Prediction

Table 5 ROC Analysis: Optimal Cutoffs for Predicting Treatment Success

Predictor	AUC	95% CI	p-value	Optimal Cutoff	Sensitivity, %	Specificity, %	PPV, %	NPV, %
Baseline ALT	0.82	0.74-0.90	<0.001	28.5 U/L	81.3	74.6	74.2	81.6
Baseline HOMA-IR	0.69	0.58-0.80	0.002	1.65	68.8	67.3	64.7	71.2
Baseline Triglycerides	0.66	0.55-0.77	0.008	135 mg/dL	64.6	63.6	62.0	66.0
Hepatic Steatosis Index	0.67	0.57-0.77	0.005	38.0	66.7	65.5	64.0	68.1

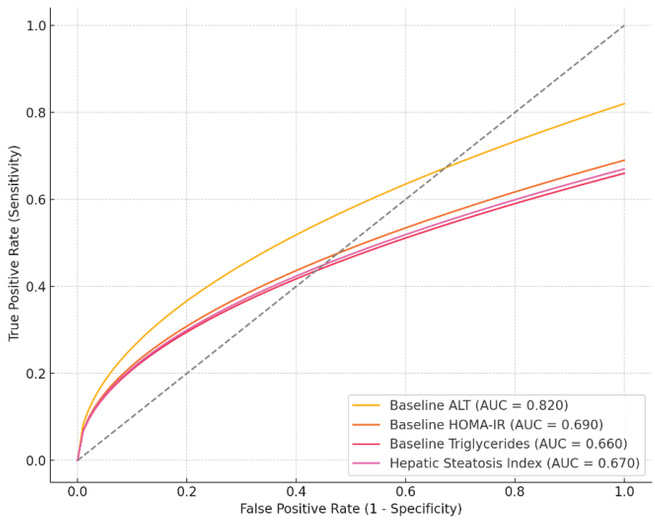


Figure 1 – ROC curve for predicting treatment success

Correlation Analysis

Table 6 Correlation Analysis: Hepatic Function Improvements and Metabolic Parameters

Variables	ALT Reduction	ALP Reduction	HSI Improvement	Weight Loss	HbA1c Reduction
ALT Reduction	1.00	0.58**	0.72***	0.64***	0.51**
ALP Reduction	0.58**	1.00	0.62***	0.52**	0.44*
HSI Improvement	0.72***	0.62***	1.00	0.71***	0.58**
Weight Loss	0.64***	0.52**	0.71***	1.00	0.62***
HbA1c Reduction	0.51**	0.44*	0.58**	0.62***	1.00
Insulin Reduction	0.48**	0.41*	0.55**	0.59**	0.76***
HOMA-IR Reduction	0.52**	0.46*	0.59**	0.61***	0.78***
Triglyceride Reduction	0.61***	0.48**	0.71***	0.67***	0.54**

**p<0.05, **p<0.01, ***p<0.001*

Discussion

This study demonstrates that a culturally adapted low-carbohydrate Paleolithic diet produces rapid hepatic function improvements in South Asian adults, with the novel contribution of predictive modeling enabling personalized intervention strategies. The development of validated predictive tools addresses a critical gap in personalized nutrition for NAFLD management.

Primary Finding: Predictive Modeling for Treatment Success

The multiple regression model explaining 64% of variance in hepatic response represents a significant advancement in personalized NAFLD management. The identification of baseline ALT (β=0.58), insulin resistance (β=0.34), and triglycerides (β=0.28) as independent predictors provides clinicians with evidence-based tools for patient selection and outcome prediction. The model's strong predictive accuracy and cross-validation stability (R^2 = 0.61) support its clinical utility.

The ROC analysis revealing baseline ALT >28.5 U/L as optimal predictor of treatment success (AUC = 0.82) provides a practical clinical threshold. This cutoff's high sensitivity (81.3%) and specificity (74.6%) enable clinicians to identify patients most likely to benefit from this intervention, optimizing resource allocation and treatment planning.

Hepatic Function Improvements: Mechanisms and Clinical Significance

The substantial hepatic improvements observed (17.1% ALT reduction, 7.4% ALP reduction, 10.7% HSI improvement) exceed those typically achieved with general weight loss interventions and suggest specific hepatoprotective effects of the low-carbohydrate approach [15]. The greater response in participants with baseline ALT >28.5 U/L (34.2% reduction) indicates that individuals with existing hepatic dysfunction derive maximum benefit.

The mechanism underlying these improvements involves multiple pathways: reduced hepatic de novo lipogenesis through carbohydrate restriction, enhanced fatty acid oxidation, decreased insulin resistance, and improved hepatic insulin sensitivity [16]. The strong correlation between ALT reduction and insulin resistance improvement ($r = 0.52$) supports the central role of insulin sensitivity in hepatic function recovery.

Novel Contribution: Population-Specific Predictive Tools

This study addresses a critical gap in South Asian NAFLD management by developing the first validated predictive model for dietary intervention success in this population. The cultural adaptation of the Paleolithic diet, combined with predictive modeling, represents a novel approach to personalized hepatic health management. The high treatment success rate (46.6%) and robust predictive accuracy demonstrate the feasibility of this approach.

The predictive model's components (baseline ALT, insulin resistance, triglycerides) reflect the underlying pathophysiology of NAFLD in South Asian populations, where insulin resistance and dyslipidemia are primary drivers of hepatic dysfunction [17]. The negative association with age ($\beta = -0.23$) suggests that younger individuals may have greater capacity for hepatic function recovery.

Clinical Implications and Personalized Medicine

The predictive tools developed in this study enable clinicians to:

1. Risk Stratification: Identify patients most likely to benefit from dietary intervention using the ALT >28.5 U/L threshold
2. Resource Optimization: Focus intensive interventions on high-probability responders
3. Patient Counseling: Provide evidence-based expectations using the regression equation
4. Treatment Planning: Tailor intervention intensity based on predicted response probability

The combined logistic regression model achieving AUC = 0.87 demonstrates that multiple predictors provide superior accuracy compared to single biomarkers, supporting the development of comprehensive predictive algorithms for clinical practice.

Integrated Metabolic Benefits

The strong correlations between hepatic improvements and metabolic parameters ($r = 0.52$-0.72) demonstrate the integrated nature of metabolic health improvement. Responders achieved superior outcomes across all metabolic parameters, suggesting that hepatic function improvement serves as a marker of comprehensive metabolic recovery.

The concurrent improvements in insulin resistance (HOMA-IR reduction in responders) and triglycerides align with known mechanisms of hepatic fat reduction through improved insulin sensitivity and decreased hepatic lipogenesis [18]. These findings support the use of hepatic function markers as surrogate endpoints for metabolic health improvement.

Statistical Robustness and Clinical Validity

The multiple regression model's strong performance ($R^2 = 0.64$) and cross-validation stability support its clinical reliability. The ROC analysis providing AUC values >0.8 for the primary predictor indicates excellent discriminative ability for clinical decision-making [19]. The model's equation enables precise probability calculations for individual patients, facilitating personalized treatment approaches.

Limitations and Future Directions

Study limitations include single-arm design, absence of histological validation, and short-term follow-up. The predictive model requires external validation in independent South Asian populations. Future research should include randomized controlled trials with histological endpoints, longer-term follow-up to assess sustainability of predictions, external validation in diverse populations, and cost-effectiveness analysis of predictive model implementation.

Conclusion

This study provides the first evidence-based predictive model for hepatic function improvement with culturally adapted dietary intervention in South Asian adults. The identification of baseline ALT >28.5 U/L as optimal predictor of treatment success, combined with the multiple regression model explaining 64% of response variance, enables personalized approaches to NAFLD management in this population.

The substantial hepatic improvements (17.1% ALT reduction, 7.4% ALP reduction) integrated with comprehensive metabolic benefits support the clinical utility of this intervention. The predictive tools developed enable clinicians to optimize patient selection, resource allocation, and treatment planning, representing a significant advancement in personalized nutrition for hepatic health.

These findings have important implications for NAFLD prevention and management in South Asian populations, providing evidence-based tools for implementing culturally adapted dietary interventions with predicted outcomes. The integration of predictive modeling with intervention effectiveness represents a novel approach to personalized hepatic health management.

Author Contributions: Conceptualization, H.V. and V.C.K.; methodology, H.V. and V.C.K.; validation, H.V. and V.C.K.; formal analysis, H.V.; investigation, H.V.; resources, V.C.K.; data curation, H.V.; writing – original draft preparation, H.V.; writing – review and editing, H.V. and V.C.K.; visualization, H.V.; supervision, V.C.K.; project administration, H.V.; funding acquisition – not applicable. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: The authors acknowledge the dietitians helped in recruiting the participants. Furthermore, the authors thanks all the survey participants for their valuable contribution.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. Teng ML, Ng CH, Huang DQ, Chan KE, Tan DJ, Tan C, Lim WH, Yang JD, Tan E, Muthiah M, Siddiqui MS, Sanyal AJ, Loomba R, Wang JW, Nah B, Huang R, Schwabe RF, Lim SG, Carey WB, Noureddin M, Lim KH, Ngoh CL, Mak LY, Yuen MF, Seto WK, Kulkarni AV, Singh SP, Dan YY, Goh GBB, Dhanasekaran R, Low HM, Leong J, Wong VWS, Singal AK, Mahady SE, Nakajima A, Lee KS, Chan WK, Targher G, Duseja A, George J, Chawla Y. Global incidence and prevalence of nonalcoholic fatty liver disease. *Clinical and Molecular Hepatology.* 2023; 29(Suppl): S32-S42. https://doi.org/10.3350/cmh.2022.0365
2. Shalimar, Elhence A, Bansal B, Gupta H, Anand A, Vishnu VY, Kalaivani M, Bhatia V, Srivastava DN, Garg PK, Datta Gupta S. Prevalence of Non-alcoholic Fatty Liver Disease in India: A Systematic Review and Meta-analysis. *Journal of Clinical and Experimental Hepatology.* 2022; 12(3): 818-829. https://doi.org/10.1016/j.jceh.2021.11.010
3. Choudhuri G, Shah S, Kulkarni A, Sachdeva S, Roy A, Dhiman RK, Venugopal K, Shalimar, Duseja A, Wadhawan M, Chawla Y, Mishra P, Mukherjee A, Saraf N, Eapen CE, Jain D, Benjamin J, Saraya A, Tandon R, Rathi P. Non-alcoholic Steatohepatitis in Asians: Current Perspectives and Future Directions. *Cureus*. 2023; 15(8): e42852. https://doi.org/10.7759/cureus.42852
4. Premkumar M, Anand AC. Lean Fatty Liver Disease: Through Thick and Thin. Journal of Clinical and Experimental Hepatology. 2021; 11(5): 523-527. https://doi.org/10.1016/j.jceh.2021.06.011
5. Mohan V, Unnikrishnan R, Shobana S, Malavika M, Anjana RM, Sudha V. Are excess carbohydrates the main link to diabetes & its complications in Asians? *Indian Journal of Medical Research.* 2018; 148(5): 531-538. https://doi.org/10.4103/ijmr.IJMR_1698_18
6. Redman LM, Heilbronn LK, Martin CK, de Jonge L, Williamson DA, Delany JP, Ravussin E. Metabolic and behavioral compensations in response to caloric restriction: implications for the maintenance of weight loss. *PLoS One.* 2009; 4(2): e4377. https://doi.org/10.1371/journal.pone.0004377
7. Kandula NR, Bernard V, Dave S, Patel M, Seguil P, Talegawkar S, Shah N, Kang A, Saucedo L, Venkatraman R, Bharucha A, Narayan KMV, Prabhakaran D, Shah NS. The South Asian Healthy Lifestyle Intervention (SAHELI) trial: Protocol for a mixed-methods, hybrid effectiveness implementation trial for reducing cardiovascular risk in South Asians in the United States. *Contemporary Clinical Trials.* 2020; 92: 105995. https://doi.org/10.1016/j.cct.2020.105995
8. Keating SE, Chawla Y, De A, George ES. Lifestyle intervention for metabolic dysfunction-associated fatty liver disease: a 24-h integrated behavior perspective. *Hepatology International.* 2024; 18(Suppl 2): 959-976. https://doi.org/10.1007/s12072-024-10663-9
9. Puri P, Dhiman RK, Taneja S, Tandon P, Valsan A, Duseja A, Eapen CE, Venkataraman J, Govil A, Gupta N, Merh N, Merh N, Chaudhary A, Singh SP, Amarapurkar DN, Benjamin J, Kalro RH. Nutrition in Chronic Liver Disease: Consensus Statement of the Indian National Association for Study of the Liver. *Journal of Clinical and Experimental Hepatology.* 2021; 11(1): 97-143. https://doi.org/10.1016/j.jceh.2020.09.003
10. Yaghoubi F, Darand M, Vasmehjani AA, Dastan S, Hajipour A, Yaghoubi S, Goodarzi-Khoigani M, Hosseinzadeh M. Adherence to low carbohydrate diets and non-alcoholic fatty liver disease: a case control study. *BMC Nutrition.* 2022; 8(1): 140. https://doi.org/10.1186/s40795-022-00625-5
11. Senn S. Mastering variation: variance components and personalised medicine. *Statistics in Medicine.* 2016; 35(7): 966-977. https://doi.org/10.1002/sim.6739
12. Sadr H, Nazari M, Khodaverdian Z, Alavi MA, Moghaddam AA, Zarrabi A, Sharifi E, Falahati M, Varma RS, Rabiee N. Unveiling the potential of artificial intelligence in revolutionizing disease diagnosis and prediction: a comprehensive review of machine learning and deep learning approaches. *European Journal of Medical Research.* 2025; 30(1): 418. https://doi.org/10.1186/s40001-025-02680-7
13. Yu HS, Jiang H, Li MK, Wu YY, Li C, Zhou YJ, Qi YN, Yang YL, Wang XJ, Jiang W, Yuan X, Pan Q, Zhang RN, Fan JG. Lowering the threshold of alanine aminotransferase for enhanced identification of significant hepatic injury in chronic hepatitis B patients. World Journal of Gastroenterology. 2023; 29(35): 5166-5177. https://doi.org/10.3748/wjg.v29.i35.5166
14. Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. *Nature Medicine*. 2018; 24(7): 908-922. https://doi.org/10.1038/s41591-018-0104-9
15. Smith GI, Shankaran M, Yoshino M, Schweitzer GG, Chondronikola M, Beals JW, Okunade AL, Patterson BW, Nyangau E, Field T, Sirlin CB, Talukdar S, Hellerstein MK, Klein S. Insulin resistance drives hepatic de novo lipogenesis in nonalcoholic fatty liver disease. *Journal of Clinical Investigation.* 2020; 130(3): 1453-1460. https://doi.org/10.1172/JCI134165
16. Verbakel JY, Steyerberg EW, Uno H, De Cock B, Wynants L, Collins GS, Van Calster B. ROC curves for clinical prediction models part 1. ROC plots showed no added value above the AUC when evaluating the performance of clinical prediction models. Journal of Clinical Epidemiology. 2020; 126: 207-216. https://doi.org/10.1016/j.jclinepi.2020.01.028
17. Albhaisi S, Chowdhury A, Sanyal AJ. Non-alcoholic fatty liver disease in lean individuals. *JHEP Reports.* 2019; 1(4): 329-341. https://doi.org/10.1016/j.jhepr.2019.08.002
18. Amare F, Kiflu A, Taddese A. Effects of concurrent continuous aerobic and short rest resistance exercise training on metabolic biomarkers in type 2 diabetes patients: a systematic review and meta-analysis. *Diabetology & Metabolic Syndrome.* 2025; 17(1): 290. https://doi.org/10.1186/s13098-025-01838-x
19. Roumeliotis S, Schurgers LJ, Tsalikakis DG, Georgianos PI, Liakopoulos V, Eleftheriadis T, Thodis E, Mayer G, Stamatelou K. ROC curve analysis: a useful statistic multi-tool in the research of nephrology. *International Urology and Nephrology.* 2024; 56(8): 2651-2658. https://doi.org/10.1007/s11255-024-04022-8

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17431

Risk Management in Gynecological Departments of Hospitals in Almaty

Bakhyt M. Sultanbekova[1], Marzhan Brimzhanova[2], Gulnara Kandygulova[2], Nurgul Alekenova[3], Anuar Akhmetzhan[4], Natalya Pirozhenko[3]

[1]Doctoral program, Al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Public health development department, Kazakhstan's Medical University «KSPH», Almaty, Kazakhstan
[3]Department of evidence based medicine, pathological anatomy, Marat Ospanov West-Kazakhstan Medical University, Aktobe, Kazakhstan
[4]Health policy and management department, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan

Received: 2025-02-05.
Accepted: 2025-11-18.

J Clin Med Kaz 2025; 22(6): 80–88

Corresponding author:
Nurgul Alekenova.
E-mail: alekenova@zkmu.kz.
ORCID: 0000-0001-5751-1678.

ABSTRACT

Introduction. Kazakhstan's healthcare system, which has implemented accreditation since 2010, emphasizes risk management, and with the rise in gynecological hospitalizations in Almaty, it is crucial to assess the preparedness of gynecological departments to ensure patient safety.

Aim: to explore risk management organization in gynecological departments of hospitals in Almaty.

Methods. In 2024, a cross-sectional survey was conducted among gynecologists and nurses in Almaty's multifunctional hospitals. The survey included 63 gynecologists and 45 nurses, with a sample size calculated based on a 95% confidence level and 5% margin of error. Data analysis was performed using chi-square tests and processed with MS Excel and SPSS, ensuring reliable results for comparing responses between the two groups.

Results. The majority of respondents (over 10 years of experience) reported that patient safety updates were provided regularly, but concerns about inadequate financial resources and infrastructure persisted, with 43.5% believing that the hospital employs risk management specialists. A third of respondents felt that the hospital's risk management efforts were insufficient, and more than half appreciated training in patient safety, though about a third felt it was inadequate. Statistically significant differences were observed, with 30.2% of doctors and 24.4% of nurses reporting errors detected before reaching the patient (p=0.006), and 59.3% of respondents reported no patient safety events in the past 12 months. Nurses rated their departments more positively than doctors, with 60% rating safety as satisfactory, while 36.5% of doctors felt management only focused on patient safety after an adverse event occurred (p<0.001).

Conclusion. This study highlights patient safety challenges in Almaty hospitals, including inadequate resources, infrastructure, and training, while emphasizing the need for better communication and improvements in safety ratings from both doctors and nurses.

Keywords: gynecology, risk management, quality of care, hospital.

Introduction

Policymakers play an important role in achieving universal health coverage (UHC) by designing and executing policies that ensure all people have access to critical health care without incurring financial hardship [1,2]. To promote excellent health outcomes, UHC demands not just increased access to healthcare but also improved service quality.

Quality of care refers to how well healthcare services offered to individuals or groups fulfill defined criteria, are effective, and enhance patient outcomes [3,4]. It has multiple characteristics, including safety, patient-centeredness, timeliness, efficiency, and equity. High-quality care entails providing the right care at the right time, with an emphasis on avoiding harm, improving the patient experience, and tailoring care to individual requirements.

A good safety culture within an organization assures service users, staff, and the public that there is a commitment to providing high-quality, safe, and effective care. Robust risk management includes a blame-free reporting culture and learning from clinical errors, as well as a proactive approach to measuring patient safety indicators [5].

In order to ensure patient safety and uphold high standards of care, risk management is crucial in hospitals. It improves overall clinical results by preventing medical errors, which can result in patient damage and liability claims. Proper risk management leads to a safer environment for both patients and workers [6,7]. However, the findings indicate that many physicians do not feel free to report errors. Research on physician decision making (e.g., the person-who effect, regret avoidance, and the availability heuristic) may provide useful insights into medical error reporting issues [8]. According to studies conducted in both developing and developed countries, variables such as supply shortages, a lack of skilled workers, delayed care, and poor communication all contribute to medical errors and adverse outcomes. These concerns, including second and third-order delays in care, are frequently caused by inadequate infrastructure, referral mechanisms, and coordination, all of which increase the chance of errors [9,10]. In wealthy countries, delays may occur as a result of fragmented healthcare systems or a lack of continuity of care [11]. Medical record reviews, while essential, may be hampered by biases and inadequate data, which fail to capture all systemic or environmental elements that contribute to errors [12]. Therefore, implementation process for open discussion is important, for instance, Northwell Health's Obstetrics and Gynecology Service Line implemented a weekly Safety Call, contributing to a 19% decrease in adverse outcomes and significant reductions in insurance premiums by fostering collaboration, identifying risks, and improving patient safety across its 10 maternity hospitals [13].

While medical errors and adverse events are frequently documented in general medicine and surgery, less is known about errors in obstetrics and gynecology. The study discovered that existing hospital reporting systems missed several errors, particularly those involving communication and cooperation in obstetrics and gynecology. Physician participation in the reporting process, which was integrated into teaching rounds, proved useful in detecting errors that were not caught by traditional approaches [14].

The healthcare system of Kazakhstan aims to improve the quality of medical care, and in this regard, the accreditation process has been in place since 2010 [15]. One of the standards of the accreditation process is risk management, which is an important element in ensuring safe care for patients [16]. In recent years, there has been an increase in hospitalization of women with gynecological diseases [17] in Almaty city. Accordingly, it is important to study how gynecological departments are prepared to ensure patient safety in existing inpatient settings. Thus, the aim of our study is to examine how risk management, particularly risk identification, mitigation strategies and monitoring practices organized in the gynecological departments of hospitals in Almaty.

Methods

A cross-sectional survey of gynecologists working in Almaty's multifunctional hospitals was undertaken in 2024. The questionnaire, developed by the authors based on existing literature, was designed to evaluate the unit's performance, the role of branch managers, and communication practices in gynecologic department. To assess content validity, the initial draft was reviewed by a panel of the three experts in gynecology and two specialists in healthcare quality and clinical audit. Additionally, validity was evaluated through a pilot test with ten gynecologists and nurses, who were not included in the final analysis. Participants were selected using purposive sampling approach based on following criteria: current employment in a gynecology department (gynecologists and nurses) of Almaty's multifunctional hospitals, and direct involvement in patient care. Participation was voluntary and anonymous, and the survey was administered online via Google Forms in collaboration with the Health Department and hospital management. This method ensured that only those with appropriate professional roles and experience contributed to the findings, thereby enhancing the relevance of the data, though it may limit generalizability beyond this specific clinical setting.

There were 67 gynecologists and 53 nurses registered in the city's hospital gynecology departments. Using a cross-sectional study design method, the required sample size was determined to be 57 gynecologists and 47 nurses, with a 95% confidence level and a 5% margin of error. To account for possibly partial responses and boost the trustworthiness of the data, we sought to include a greater number of respondents, resulting in the gathering of 63 completed questionnaires from gynecologists and 45 from nurses.

We used the chi-square (χ^2) test to compare replies between gynecologists and nurses. This test examines categorical variables: socio-demographic variables (profession and work experience) and 40 core variables, related to risk management structures, patient safety practices, reporting behavior, and perceptions of hospital management. We determined statistical significance by comparing the generated χ^2 statistic to the crucial value. If the statistic surpassed the critical value, we rejected the null hypothesis, indicating a significant difference.

All statistical analyses were carried out with MS Excel and SPSS13, ensuring accurate and trustworthy data processing. The Local Committee on Bioethics in Kazakhstan (IRB-A832, 21 May 2024) extensively reviewed and approved the study design, verifying that it met ethical standards and procedures for conducting research involving human participants.

Results

A total of 108 healthcare professionals participated in the study, including 63 gynecologists and 45 nurses working in gynecology departments of multifunctional hospitals in Almaty. The majority of respondents (63.0%) had more than 11 years of professional experience, with 65.1% of gynecologists and 60.0% of nurses falling into this category.

Approximately half of the participants in both groups stated that responsible persons have provided regular information regarding patient safety actions during the previous 12 months, and that the coordinator's role in promoting patient safety is primarily focused on risk management activities. Notably, 43.5% of respondents, including the highest proportion of nurses (51.1%), believed that the hospital employs specialists with a formal workload dedicated to risk management operations, while 46% of doctors and 40% of nurses disagreed. More over one-third of respondents in both categories expressed worry that the hospital does not commit enough financial resources to promote patient safety measures and lacks the physical infrastructure to assure safety. On a positive note, more than half of respondents praised the hospital's efforts to train its employees in areas such as risk management, quality management, and patient safety,

Table 1 Risk Management in Hospital

Questions		Gynecologists - n (%)	Nurse - n (%)	Nurse - n (%)	Total
Work experience	Less than 1 year	**n (%)**	**P value**		0.807
	1 to 5 years	8 (12.7)	7(15.6)	15(13.9)	
	6 to 10 years	12 (19.0)	8(17.8)	20(18.5)	
	11 or more years	41 (65.1)	27(60.0)	68(63.0)	
	Total	63 (100.0)	45(100.0)	108(100.0)	
Have you been provided with a progress report on patient safety in the last 12 months	Yes, clear enough;	29 (47.5)	26(57.8)	55(51.9)	0.018
	Yes, but not clear enough	8 (13.1)	12(26.7)	20(18.9)	
	No	24 (39.3)	7(15.6)	31(29.2)	
	Total	61 (100.0)	45(100.0)	106(100.0)	
Does the hospital have a formal workload dedicated to risk management activities	yes	24 (38.1)	23(51.1)	47(43.5)	0.328
	no	29 (46.0)	18(40.0)	47(43.5)	
	other	10 (15.9)	4(8.9)	14(13.0)	
	Auditor probably	1 (1.6)		1(0.9)	
	Total	64	45	109	
Is the patient safety improvement coordinator dedicated exclusively to risk management activities	yes	29 (46.0)	29(64.4)	58(53.7)	0.144
	no	8 (12.7)	5(11.1)	13(12.0)	
	We do not have such a specialist	26(41.3)	11(24.4)	37(34.3)	
	Total	63 (100.0)	45(100.0)	108	
Are financial resources allocated to promote patient safety activities	Yes, but not enough	20(31.7)	15(33.3)	35(32.4)	0,217
	Yes	19(30.2)	20(44.4)	39(36.1)	
	No	15(23.8)	8(17.8)	23(21.3)	
	I don't know	9(14.3)	2(4.4)	11(10.2)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital have adequate physical infrastructure to ensure patient safety	Yes, but not enough	25(39.7)	18(40.0)	43(39.8)	0.923
	Yes	31(49.2)	23(51.1)	54(50.0)	
	No	4(6.3)	3(6.7)	7(6.5)	
	I don't know	3(4.8)	1(2.2)	4(3.7)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital provide sufficient resources to implement risk management activities	Yes, but not enough	26(41.3)	21(46.7)	47(43.5)	0.612
	Yes	30(47.6)	22(48.9)	52(48.1)	
	No	2(3.2)	1(2.2)	3(2.8)	
	I don't know	5(7.9)	1(2.2)	6(5.6)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Has the hospital facilitated the training of its staff in this area (risk management, quality management, patient safety, etc.)	Yes, but not enough;	19(30.2)	13(28.9)	32(29.6)	0.955
	Yes, quite clearly	32(50.8)	25(55.6)	57(52.8)	
	No	7(11.1)	5(11.1)	12(11.1)	
	I don't know	3(4.8)	1(2.2)	4(3.7)	
	At my own expense, tuition	2(3.2)	1(2.2)	3(2.8)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Has a safety culture assessment been conducted in the last 12 months	Yes, but not enough	18(28.6)	11(24.4)	29(26.9)	0.260
	Yes	30(47.6)	29(64.4)	59(54.6)	
	No	8(12.7)	2(4.4)	10(9.3)	
	I don't know	7(11.1)	3(6.7)	10(9.3)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Have the results of the safety culture assessment been communicated to clinical. administrative and care staff	Yes. but not enough	15(23.8)	8(17.8)	23(21.3)	0.288
	Yes	30(47.6)	29(64.4)	59(54.6)	
	No	10(15.9)	3(6.7)	13(12.0)	
	I don't know	7(11.1)	3(6.7)	10(9.3)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Following the safety culture assessment. have any interventions been implemented to improve the findings identified in the assessment	Yes. but not enough	17(27.0)	14(31.1)	31(28.7)	0.552
	Yes	31(49.2)	26(57.8)	57(52.8)	
	No	8(12.7)	2(4.4)	10(9.3)	
	I don't know	5(7.9)	2(4.4)	7(6.5)	
	Lack of funding. But there is a plan	2(3.2)	1(2.2)	3(2.8)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	

Does the hospital use an internal incident reporting system	Yes. but not enough	16(25.4)	10(22.2)	26(24.1)	0.866
	Yes	44(69.8)	32(71.1)	76(70.4)	
	No	3(4.8)	3(6.7)	6(5.6)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital monitor compliance with international patient safety targets	Yes. but not enough	17(27.0)	9(20.0)	26(24.1)	0.471
	Yes	38(60.3)	29(64.4)	67(62.0)	
	No	3(4.8)	5(11.1)	8(7.4)	
	I don't know	5(7.9)	2(4.4)	7(6.5)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital use outcome indicators to identify risks	Yes. but not enough	6(9.5)	2(4.4)	8(7.4)	0.316
	Yes	44(69.8)	38(84.4)	82(75.9)	
	No	8(12.7)	4(8.9)	12(11.1)	
	I don't know	5(7.9)	1(2.2)	6(5.6)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital have a death review committee that has met in the last six months (minutes)	Yes. but not enough	12(19.0)	10(22.2)	22(20.4)	0.870
	Yes	44(69.8)	32(71.1)	76(70.4)	
	No	5(7.9)	2(4.4)	7(6.5)	
	I don't know	2(3.2)	1(2.2)	3(2.8)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Is direct observation used to identify risks? (e.g. hand hygiene checks. contact precautions for isolated patients. protective barriers when introducing central	Yes. but not enough	10(15.9)	7(15.6)	17(15.7)	0.435
	Yes	48(76.2)	37(82.2)	85(78.7)	
	No	5(7.9)	1(2.2)	6(5.6)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the security staff conduct walk-throughs of areas to identify risks	Yes. but not enough	13(20.6)	7(15.6)	20(18.5)	0.296
	Yes	39(61.9)	35(77.8)	74(68.5)	
	No	4(6.3)	1(2.2)	5(4.6)	
	I don't know	7(11.1)	2(4.4)	9(8.3)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital use tools for qualitative analysis of causes and contributing factors for patient safety (flow chart. cause and effect diagram. brainstorming. etc.)	Yes. but not enough	7(11.1)	8(17.8)	15(13.9)	0.350
	Yes	38(60.3)	20(44.4)	58(53.7)	
	No	5(7.9%)	3(6.7%)	8(7.4%)	
	I don't know	13(20.6%)	14(31.1%)	27(25.0%)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital use tools for quantitative analysis of causes or contributing factors of risks (histogram. stratification. Pareto chart. and control chart)	Yes. but not enough	7(11.1%)	10(22.2)	17(15.7)	0.273
	Yes	34(54.0%)	18(40.0)	52(48.1)	
	No	1(1.6)	2(4.4)	3(2.8)	
	I don't know	21(33.3)	15(33.3)	36(33.3)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital use any risk prioritization matrix based on criteria of severity and frequency of events. Affecting patient safety	Yes. but not enough	7(11.1)	7(15.6)	14(13.0)	0.633
	Yes	27(42.9)	20(44.4)	47(43.5)	
	No	2(3.2)	3(6.7)	5(4.6)	
	I don't know	27(42.9)	15(33.3)	42(38.9)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the quality specialist assess the adequacy of measures to control or reduce patient or other risks	Yes. but not enough	11(17.5)	9(20.0)	20(18.5)	0.286
	Yes	32(50.8)	24(53.3)	56(51.9)	
	No		2(4.4)	2(1.9)	
	I don't know	20(31.7)	10(22.2)	30(27.8)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Has the hospital implemented key clinical protocols for patient safety	Yes. but not enough	8(12.7)	8(17.8)	16(14.8)	0.865
	Yes	40(63.5)	27(60.0)	67(62.0)	
	No	2(3.2)	2(4.4)	4(3.7)	
	I don't know	13(20.6)	8(17.8)	21(19.4)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Has the hospital implemented action plans in response to adverse events investigated	Yes. but not enough	5(7.9)	8(17.8)	13(12.0)	0.308
	Yes	38(60.3)	25(55.6)	63(58.3)	
	No	1(1.6)	2(4.4)	3(2.8)	
	I don't know	19(30.2)	10(22.2)	29(26.9)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	

Does the hospital describe the person responsible for implementing risk mitigation activities	Yes. but not enough	3(4.8)	9(20.0)	12(11.1)	0.058
	Yes	35(55.6)	23(51.1)	58(53.7)	
	No	2(3.2)		2(1.9)	
	I don't know	23(36.5)	13(28.9)	36(33.3)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital describe a schedule for implementing risk mitigation activities	Yes. but not enough	7(11.1)	8(17.8)	15(13.9)	0.299
	Yes	33(52.4)	27(60.0)	60(55.6)	
	No	2(3.2)		2(1.9)	
	I don't know	21(33.3)	10(22.2)	31(28.7)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital describe and measure the implementation and effectiveness of risk mitigation activities	Yes. but not enough	8(12.7)	8(17.8)	16(14.8)	0.508
	Yes	29(46.0)	24(53.3)	53(49.1)	
	No	1(1.6)		1(.9)	
	I don't know	25(39.7)	13(28.9)	38(35.2)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Does the hospital describe the resources needed for risk mitigation activities in the delivery of care	Yes. but not enough	13(20.6)	13(28.9)	26(24.1)	0.539
	Yes	30(47.6)	21(46.7)	51(47.2)	
	I don't know	20(31.7)	11(24.4)	31(28.7)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Is reporting of adverse events (open disclosure of errors) standardized to patients through any institutional norms. protocol. or policy	Yes. but not enough	16(25.4)	15(33.3)	31(28.7)	0.808
	Yes	26(41.3)	18(40.0)	44(40.7)	
	No	5(7.9)	3(6.7)	8(7.4)	
	I don't know	16(25.4)	9(20.0)	25(23.1)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Has the hospital completed the full risk management cycle (identification. analysis. assessment. risk mitigation. and risk monitoring) in the past 12 months	Yes. but not enough	12(19.0)	14(31.1)	26(24.1)	0.270
	Yes	27(42.9)	19(42.2)	46(42.6)	
	I don't know	24(38.1)	12(26.7)	36(33.3)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
	Yes	27(42.9)	19(42.2)	46(42.6)	
	I don't know	24(38.1)	12(26.7)	36(33.3)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	

Table 2 Patient Safety Assessment

Questions		Gynecologists - n (%)	Nurse – n (%)	Total – n (%)	P value
When an error is detected and corrected before it reaches a patient. how often is it reported	Never	15 (23.8)	10 (22.2)	25 (23.1)	0.006
	Rarely	21 (33.3)	6 (13.3)	27 (25.0)	
	Sometimes	4 (6.3)	11 (24.4)	15 (13.9)	
	Most of the time		4 (8.9)	4 (3.7)	
	Always	19 (30.2)	11 (24.4)	30 (27.8)	
	Non-applicable or I don't know	4 (6.3)	3 (6.7)	7 (6.5)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
When an error reaches a patient and could have caused harm but does not. how often is it reported	Never	20 (31.7)	3 (6.7)	23 (21.3)	0.002
	Rarely	8 (12.7)	10 (22.2)	18 (16.7)	
	Sometimes		3 (6.7)	3 (2.8)	
	Most of the time	8 (12.7)	14 (31.1)	22 (20.4)	
	Always	23 (36.5)	12 (26.7)	35 (32.4)	
	Non-applicable or I don't know	4 (6.3)	3 (6.7)	7 (6.5)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
How many patient safety events have you reported in the last 12 months	none	44 (69.8)	20 (44.4)	64 (59.3)	0.001
	1 to 2	7 (11.1)	15 (33.3)	22 (20.4)	
	3 to 5	4 (6.3)	10 (22.2)	14 (13.0)	
	6 to 10	4 (6.3)		4 (3.7)	
	11 or more	4 (6.3)		4 (3.7)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
How would you rate your department/work area for patient safety	Bad		3 (6.7)	3 (2.8)	0.001
	Satisfactory	17 (27.0)	27 (60.0)	44 (40.7)	
	Good	19 (30.2)		19 (17.6)	
	Very good	15 (23.8)	15 (33.3)	30 (27.8)	
	Excellent	12 (19.0)		12 (11.1)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	

Table 3 Hospital Management's Actions on Risk Management

Questions		Gynecologists– n (%)	Nurse – n (%)	Total – n (%)	P value
The hospital's management's actions show that patient safety is a top priority	Strongly disagree	12(19.0)	3(6.7)	15(13.9)	0.005
	Disagree	4(6.3)	11(24.4)	15(13.9)	
	Neither agree nor disagree	8(12.7)	3(6.7)	11(10.2)	
	Agree	31(49.2)	17(37.8)	48(44.4)	
	Completely agree	8(12.7)	7(15.6)	15(13.9)	
	Not applicable or don't know		4(8.9)	4(3.7)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Hospital management provides sufficient resources to improve patient safety	Strongly disagree	12(19.0)	10(22.2)	22(20.4)	0.079
	Disagree	4(6.3)	4(8.9)	8(7.4)	
	Neither agree nor disagree	8(12.7)	3(6.7)	11(10.2)	
	Agree	27(42.9)	21(46.7)	48(44.4)	
	Completely agree	12(19.0)	3(6.7)	15(13.9)	
	Not applicable or don't know		4(8.9)	4(3.7)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Hospital management seems to be interested in patient safety only after an adverse event has occurred	Strongly disagree	16(25.4)	3(6.7)	19(17.6)	0.001
	Disagree	16(25.4)	14(31.1)	30(27.8)	
	Neither agree nor disagree	4(6.3)	7(15.6)	11(10.2)	
	Agree	23(36.5)	3(6.7)	26(24.1)	
	Completely agree	4(6.3)	3(6.7)	7(6.5)	
	Not applicable or don't know		15(33.3)	15(13.9)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Crucial information is often missed when patients are transferred from one department to another	Strongly disagree	16(25.4)	14(31.1)	30(27.8)	0.003
	Disagree	19(30.2)	7(15.6)	26(24.1)	
	Neither agree nor disagree	4(6.3)	6(13.3)	10(9.3)	
	Agree	12(19.0)		12(11.1)	
	Completely agree	4(6.3)	3(6.7)	7(6.5)	
	Not applicable or don't know	8(12.7)	15(33.3)	23(21.3)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Crucial information about patient care is often missed during shift changes	Strongly disagree	24(38.1)	11(24.4)	35(32.4)	0.001
	Disagree	15(23.8)	11(24.4)	26(24.1)	
	Neither agree nor disagree	4(6.3)	6(13.3)	10(9.3)	
	Agree	12(19.0)		12(11.1)	
	Completely agree	4(6.3)	13(28.9)	17(15.7)	
	Not applicable or don't know	4(6.3)	4(8.9)	8(7.4)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
There is ample time during shift changes to exchange all key patient care information	Strongly disagree	16(25.4)	7(15.6)	23(21.3)	0.013
	Disagree	16(25.4)	7(15.6)	23(21.3)	
	Neither agree nor disagree		7 (15.6)	7 (6.5)	
	Agree	15 (23.8)	7 (15.6)	22 (20.4)	
	Completely agree	12 (19.0)	13 (28.9)	25 (23.1)	
	Not applicable or don't know	4 (6.3)	4 (8.9)	8 (7.4)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	
Your comments	The building is old. the elevator is terrible. there are no conditions in the hospital for the staff		4 (8.9)	4 (3.7)	0.010
	Bring joint liability to the patients themselves	4 (6.3)		4 (3.7)	
	Total	63 (100.0)	45(100.0)	108 (100.0)	

albeit roughly one-third felt that the training was insufficient. Overall, over half of respondents thought risk management in the hospital was obvious and adequate, while about a third thought it was insufficient (Table 1). It is crucial to note, however, that the responses of the two groups of respondents did not differ statistically significantly.

It is statistically significant that about a third of doctors (30.2%) and nurses (24.4%) report errors that are detected and corrected before they reach the patient p=0.006. While a greater number of doctors (36.55) always and most of the time nurses (31.1%) report errors that reach the patient and could have harmed him/her, but did not p=0.002. The greatest number of respondents in both groups did not report patient safety events, 59.3% reported them in the last 12 months, while the remaining 20.4% reported 1-2 incidents p<0.001. The greatest number of nurses rated their department/work area in terms of patient safety as satisfactory (60.0%), while about a third of doctors rated it as satisfactory (27.0%) and good (30.2%) p<0.001 (Table 2). The answer to the question about the hospital management actions showing that patient safety is the top priority was mixed, with about half agreeing with this statement, while the rest were not entirely in agreement (p=0.005). More than a third of doctors (36.5%) agreed that the hospital management seemed interested in-patient safety only after an adverse event occurred, while 33.3% of nurses said they did not know this question (p<0.001). The largest number of respondents in both groups (52.5%) disagreed with the statement that important information is often missed when transferring patients from one department to another (p=0.003), as well as important information about patient care is often missed during shift changes 56.5% (p<0.001), Table 3.

Additional wishes among doctors respondents noted the importance of joint responsibility of the patients themselves 6.3%, while among nurses the old building and lack of working conditions in the hospital 8.9% (Table 3).

Discussion

Perceptions of patient safety among healthcare workers can be measured by measuring clinical staff attitudes and beliefs in order to raise awareness and implement procedures that will promote patient safety culture [18,19]. In our survey, both physicians and nurses felt that training was adequate, albeit around one-third stated that more training was required. As a result, it is critical for gynaecology department managers to identify future training requirements to increase patient safety and risk management. In addition, fostering an institutional patient safety culture is critical for reducing burnout and improving work-life balance for all hospital employees, making it a worthwhile investment for medical institutions. Initiatives include increasing staffing, providing training, improving communication, and offering employee support programs [20].

Patient safety culture in hospitals in the Netherlands, Taiwan, and the United States revealed that teamwork within departments was high in all three countries, but handovers and transitions were noted as a weakness [21]. In our study, information transfer from management to personnel was also inadequate, whereas handovers between shifts were regarded highly by respondents.

We found that a third of doctors and nurses reported errors detected and corrected before reaching the patient, while a larger proportion reported errors that reached the patient but did not cause harm, with significant differences in safety event reporting and department safety ratings between doctors and nurses. These differences warrant further exploration to understand what factors contribute to discrepancies in reporting practices. Potential influences may include differences in professional roles, perceptions of responsibility, or varying levels of confidence in institutional support following reporting. Understanding these dynamics could help tailor interventions that foster open communication, improve trust in reporting systems, and encourage consistent safety practices across professional groups [22,23].

Establishing patient safety reporting systems is critical for improving treatment because it allows healthcare institutions to gather, evaluate, and share data on incidents such as adverse events, near misses, and medical errors [24]. Hospital management's prioritization of patient safety was viewed positively by half of the respondents, with many physicians believing management only addresses safety after an adverse event, while a significant portion of nurses were uncertain. However, effective risk management in healthcare requires a top-down, interprofessional team approach, with managers playing a crucial role in implementing policies, assessing non-clinical risks, and ensuring compliance with legal standards. A comprehensive framework, based on the ISO31000 model, guides managers in conducting risk assessments and using appropriate tools to enhance patient safety and minimize potential harm. By adopting these risk management techniques, healthcare organizations can improve service effectiveness, prevent adverse events, and ensure legal compliance, thereby reducing the need for compensation and insurance claims [25-27]. Improving patient safety reporting systems is essential to address gaps identified in the study. A key measure is establishing a blame-free, non-punitive culture where healthcare workers feel safe reporting errors. This can be achieved by ensuring anonymity, clear communication channels, and focusing on learning from mistakes rather than assigning blame. Hospitals should provide regular training on the value of reporting and involve leadership in creating a transparent environment. Management must acknowledge reports, take prompt action, and provide feedback to those who submit them. A supportive reporting system helps prevent future errors, builds trust, and fosters continuous improvement. Additionally, insights from these reports can lead to systemic changes, improving safety outcomes by identifying trends and implementing corrective actions [28,29].

Strength and limitation of the study. This study's strengths include a purposive sampling approach that ensured inclusion of all eligible gynecologists and nurses directly involved in gynecological care in Almaty's multifunctional hospitals, enhancing the relevance and applicability of the findings within this clinical setting. The questionnaire was carefully developed based on existing literature, reviewed by experts, and pilot tested. The achieved sample size exceeded the calculated minimum, increasing the robustness of the data. Inclusion of both gynecologists and nurses provided a comprehensive view of risk management practices from different professional perspectives. However, the study has several limitations. First, while the survey was anonymous, self-reported data may be influenced by social desirability or recollection biases, which could impair the accuracy of responses. Second, the study was done in Almaty hospitals, which may not be representative of the entire country, limiting the findings' applicability to other parts of Kazakhstan or similar healthcare systems. Third, the cross-sectional methodology collects data at a particular point in time, restricting the investigation of long-term trends or changes in patient safety

procedures and risk management techniques. Finally, the study focused on healthcare professionals' perceptions and self-reported experiences, which may not accurately reflect actual practices or the effectiveness of risk management techniques in the hospital setting.

Future steps: To translate these findings into real-world impact, results should be shared with key stakeholders—such as hospital administrators, policymakers, and quality assurance experts—through policy briefs, professional forums, and collaborative workshops. Quality improvement specialists play a vital role in turning these insights into practice by working with clinical staff to design and implement targeted interventions. Based on our results, the top actionable recommendations include: (1) enhancing training programs tailored to identified gaps in patient safety knowledge and communication practices; (2) establishing robust, anonymous error reporting systems to promote a culture of transparency and continuous improvement; and (3) allocating financial and human resources to support staffing needs and implement institutional risk management frameworks. These steps are essential to strengthen patient safety culture and reduce preventable harm in gynecology departments and beyond. Additionally, future research could use longitudinal designs to evaluate changes in patient safety practices and risk management techniques over time, providing information about the effectiveness of implemented measures. Expanding the study to include hospitals outside of Almaty would provide a more complete picture of risk management in Kazakhstan's gynecological departments and other healthcare settings. Furthermore, conducting in-depth qualitative research with healthcare workers, such as interviews or focus groups, could elicit the underlying causes of their replies and provide more extensive insights into the problems and barriers to successful risk management. Finally, adding patient perspectives through patient-centered studies would aid in assessing their experiences with patient safety measures, so contributing to improvements in treatment from both provider and patient perspectives.

Conclusion

This study identifies both strengths and gaps in patient safety and risk management in Almaty gynecology departments, with concerns around training, infrastructure, and communication. Despite some early error detection, many incidents still reach patients, highlighting the need for systemic improvements. Sharing findings with stakeholders through policy briefs and workshops can help drive change. Quality experts should lead efforts to implement targeted training, establish anonymous reporting systems, and allocate adequate resources. These actions are critical to building a stronger safety culture and reducing preventable harm.

Author Contributions: Conceptualization, B. S.; methodology,B. S.,M.B and N. A.; validation, B. S., M. A. N.A.and A. A.; formal analysis, B.S., G. K.; investigation, B. S., M. B.,N.A and N. P.; resources, B. S. A.A. G.K. and N.P; data curation, B.S., A. A.; writing – original draft preparation, B. S. ,B. M. and N. A.; writing – review and editing, B. S. and N. A.; visualization, B. S., A.A, G.K and N.P; supervision, B. S., N.A; project administration, B. S.; funding acquisition, B. S. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. Ho CJ, Khalid H, Skead K, Wong J. The politics of universal health coverage. *Lancet*. 2022;399(10340):2066-2074. https://doi.org/10.1016/s0140-6736(22)00585-2
2. Ranabhat CL, Acharya SP, Adhikari C, Kim CB. Universal health coverage evolution, ongoing trend, and future challenge: A conceptual and historical policy review. Front Public Health. 2023;11:1041459. https://doi.org/10.3389/fpubh.2023.1041459
3. Campbell SM, Roland MO, Buetow SA. Defining quality of care. *Soc Sci Med.* 2000;51(11):1611-25. https://doi.org/10.1016/s0277-9536(00)00057-5
4. Hanefeld J, Powell-Jackson T, Balabanova D. Understanding and measuring quality of care: dealing with complexity. Bull World Health Organ. 2017;95(5):368-374. https://doi.org/10.2471/blt.16.179309
5. Holland Cathrine. Risk management in gynaecology. *Obstetrics, Gynaecology and Reproductive Medicine.*2023;33(1) 20 – 28. https://doi.org/10.1016/j.ogrm.2022.10.004
6. Klein VR. Risk Management in Obstetrics and Gynecology. *Clin Obstet Gynecol.* 2019;62(3):550-559. https://doi.org/10.1097/grf.0000000000000473
7. Carroll R, Nakamura PLB. Risk Management Handbook for Healthcare Organizations. 6th ed. San Francisco, CA: *Jossey Bass;* 2011.30 c.

8. Anderson B, Stumpf PG, Schulkin J. Medical error reporting, patient safety, and the physician. *J Patient Saf.* 2009;5(3):176-9. https://doi.org/10.1097/pts.0b013e3181b320b0
9. Aikpitanyi J, Ohenhen V, Ugbodaga P, Ojemhen B, Omo-Omorodion BI, Ntoimo L.F, Imongan W, Balogun J, Okonofua FE. Maternal death review and surveillance: The case of Central Hospital, Benin City, Nigeria. *PLoS ONE.* 2019;14:e0226075. https://doi.org/10.1371/journal.pone.0226075
10. Liu J, Liu P, Gong X, Liang F. Relating Medical Errors to Medical Specialties: A Mixed Analysis Based on Litigation Documents and Qualitative Data. Risk Manag. Health Policy. 2020;13:335–345. https:// Review of 128 quality of care mechanisms: A framework and mapping for health system stewards. *Health Policy.* 2019;124:12–24. https://doi.org/10.1016/j.healthpol.2019.11.006
11. Klemann D, Rijkx M, Mertens H, van Merode F, Klein D. Causes for Medical Errors in Obstetrics and Gynaecology. *Healthcare (Basel).* 2023;11(11):1636. https://doi.org/10.3390/healthcare11111636
12. Combs A, Klein VR. A call for safety: Anticipating and mitigating risk across an obstetrics and gynecology service line. *J Healthc Risk Manag.* 2023; 43(1):38-42. https://doi.org/10.1002/jhrm.21538
13. November M, Chie L, Weingart SN. Physician-Reported Adverse Events and Medical Errors in Obstetrics and Gynecology. Advances in Patient Safety: New Directions and Alternative Approaches. 2008; 1.
14. AlAssaf, A & AKGN, H. Healthcare Accreditation in Kazakhstan: Methods and Impact. *Health Care Academician Journal.* 2016; 3: 57. http://dx.doi.org/10.5455/sad.13-1467268550
15. Syzdykova A, Turgambayeva A, Karibekov T. Risk management in healthcare system. *Journal of Clinical Medicine of Kazakhstan.* 2014;3(33):13-6.
16. Fayzullina K.M., Sultanbekova B.M., Akhmetzhan A.D., Akhtayeva N., Khamidullina Z.G., Tazhibaeva K.N., Zhorabek S.B. Analysis of the structure of hospitalization for gynecological disease in Almaty city. *Nauka i Zdravookhranenie.* 2024; 26(1): pp. 71-77.https://newjournal.ssmu.kz/publication/508/2024-1-71-77/10.34689/SH.2024.26.1.009
17. Alahmadi HA. Assessment of patient safety culture in Saudi Arabian hospitals. *Qual Saf Health Care.* 2010;19. https://doi.org/10.1136/qshc.2009.033258
18. Hameed S, Humayun A, Yaqoob M, Rehm MH. Patient Safety Culture: A Healthcare Provider's Prospect. *Cureus*. 2023 Dec 5;15(12):e49989. https://doi.org/10.7759/cureus.49989
19. Lu L, Ko YM, Chen HY, Chueh JW, Chen PY, Cooper CL. Patient Safety and Staff Well-Being: Organizational Culture as a Resource. *Int J Environ Res Public Health.* 2022;19(6):3722. https://doi.org/10.3390/ijerph19063722
20. Wagner C, Smits M, Sorra J, Huang CC. Assessing patient safety culture in hospitals across countries. *Int J Qual Health Care.* 2013;25(3):213-21. https://doi.org/10.1093/intqhc/mzt024.
21. Godsey JA, Houghton DM, Hayes T. Registered nurse perceptions of factors contributing to the inconsistent brand image of the nursing profession. Nurs Outlook. 2020 Nov-Dec;68(6):808-821. https://doi.org/10.1016/j.outlook.2020.06.005. https://www.sciencedirect.com/science/article/pii/S2352013221000818
22. Sheikhtaheri A. Near Misses and Their Importance for Improving Patient Safety. *Iran J Public Health.* 2014;43(6):853-4. https://pubmed.ncbi.nlm.nih.gov/26110160/
23. McGowan J, Wojahn A, Nicolini JR. Risk Management Event Evaluation and Responsibilities. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK559326/
24. Ferdosi M, Rezayatmand R, Molavi Taleghani Y. Risk Management in Executive Levels of Healthcare Organizations: Insights from a Scoping Review. *Risk Manag Healthc Policy.* 2020;13:215-243. https://doi.org/10.2147/rmhp.s231712
25. Guerra F, Guzzo AS, La Rosa P, Miatto I, Vergati D, Miredi F, Gasparrini N, Guerra F, Raimondi L, Corridore D, Nardi GM, Mazur M, La Torre G, Ottolenghi L. Risk management and Healthcare responsibility. How to guarantee legal protection in Medicine. *Clin Ter.* 2021;171(1):e63-e66 https://doi.org/10.7417/ct.2021.2285
26. Mistri IU, Badge A, Shahu S. Enhancing Patient Safety Culture in Hospitals. *Cureus*. 2023 Dec 27;15(12):e51159. https://doi.org/10.7759/cureus.51159.
27. van Baarle E, Hartman L, Rooijakkers S et al. Fostering a just culture in healthcare organizations: experiences in practicedoi.org/10.2147/RMHP.S246452
28. Tello JE., Barbazza E, Waddell K. *BMC Health Serv Res.* 22, 1035 (2022). https://doi.org/10.1186/s12913-022-08418-z

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17449

Classification of Combined MLO and CC Mammographic Views Using Vision–Language Models

Beibit Abdikenov[1], Nurbek Saidnassim[1], Birzhan Ayanbayev[1], Aruzhan Imasheva[1]

[1]Science and Innovation Center "Artificial Intelligence", Astana IT University, Astana, Kazakhstan

Received: 2025-08-16.
Accepted: 2025-11-21.

J Clin Med Kaz 2025; 22(6):89–95

Corresponding author:
Beibit Abdikenov.
E-mail: beibit.abdikenov@astanait.edu.kz.
ORCID: 0009-0004-3790-2339.

ABSTRACT

Background: Breast cancer remains one of the leading causes of cancer-related deaths among women globally. Early detection through mammographic screening significantly improves survival rates, but the interpretation of mammograms is time-consuming and requires extensive expertise.

Methods: We utilized six publicly available datasets, preprocessing paired craniocaudal (CC) and mediolateral oblique (MLO) views into dual-view concatenated images. Three vision-language models (VLMs)—Quantized Qwen2-VL-2B, Quantized SmolVLM (Idefics3-based), and MammoCLIP—were evaluated using two adaptation strategies: full supervised fine-tuning (SFT) and Linear Probing (LP). EfficientNet-B4 served as a CNN baseline.

Results: Experiments show that while EfficientNet-B4 achieved the highest F1-score (0.5810), VLMs delivered competitive results with additional report generation capability. MammoCLIP exhibited the best VLM performance (F1 = 0.4755, ROC-AUC = 0.6906) under LP, outperforming general-purpose VLMs, which struggled with recall despite high precision. SmolVLM demonstrated balanced performance under full fine-tuning (F1 = 0.5101, ROC-AUC = 0.6304), indicating strong adaptability in resource-efficient setups.

Conclusion: These findings highlight that domain-specific pretraining significantly enhances VLM effectiveness in mammography classification. Beyond classification, VLMs enable structured reporting and interactive decision support, offering promising avenues for clinical integration despite slightly lower predictive performance compared to specialized CNNs.

Keywords: Deep learning, vision-language models, mammography, breast cancer, medical imaging, AI in Radiology.

Introduction

Breast cancer is the most diagnosed cancer and a major cause of mortality among women worldwide [1]. Mammographic screening is a key tool for early detection, and its four standard views—L-CC, L-MLO, R-CC, and R-MLO—are typically assessed using the BI-RADS system [3]. Increasing screening demand and the shortage of expert radiologists highlights the need for AI systems that can support mammogram interpretation and report generation.

Deep learning has already shown strong potential for breast cancer diagnosis [4]. More recently, vision-language models (VLMs) have emerged as powerful tools capable of integrating image and text information. Although VLMs are trained on natural image–text pairs, studies suggest they can transfer effectively to medical imaging tasks [5–9].

In this work, we evaluate three VLMs for abnormality classification and report generation using multiview mammograms. Our contributions are:

1. Training on six heterogeneous datasets to improve generalization across imaging devices.
2. Using quantized model parameters to reduce memory usage and accelerate inference.

Related work

Originally developed for general image-text caption alignment tasks, VLMs such as CLIP [10] and MedCLIP [11] and Mammo-CLIP [12] have been increasingly adapted to medical imaging, where they show potential for robust image classification even with limited labeled data [13]. The development of medically oriented models has further improved the efficiency of clinical applications, for example the MammoVLM [14], which is adapted for mammography, improves diagnostic efficiency through carefully tailored prompts and multimodal pre-training.

Evaluations across a wide range of medical imaging tasks—including skin lesion analysis, blood cell detection, ultrasound, chest X-ray interpretation, and MRI anomaly identifications show that vision-language models consistently perform well even when applied in zero-shot or weakly supervised settings [5]. This demonstrates a key practical advantage: VLMs can adapt to new clinical tasks or imaging modalities with minimal task-specific data, reducing the need for extensive retraining and accelerating deployment.

Recent works explore different approaches in optimizing and adapting VLMs for breast cancer imaging. Zero-shot learning has become key to handling domain shifts across heterogeneous mammography datasets. Yan [15] demonstrates how adapting CLIP for multi-distribution classification in mammography yields improved robustness under zero-shot settings. This aligns with findings by Vo et al. [16], who show that frozen VLMs like CLIP and its derivatives act as cornerstone for multimodal breast cancer prediction, even without task-specific fine-tuning.

While frozen large-scale VLMs show strong performance, several limitations affect their direct implementation in clinical practice. One of the concerns is the limited adaptability of general-purpose VLMs to highly specialized clinical tasks. For instance, LLaVA-Ultra [17], while optimized for Chinese ultrasound interpretation, shows multimodal alignment suffers from language-specific contexts, requiring substantial prompt tuning and data curation. Performance inconsistency across downstream tasks is also seen in Molina-Román et al. [18], they compare VLMs against ConvNeXt in breast density assessment and find that VLMs are outperformed by tailored vision only models, particularly when data quality is suboptimal. It suggests that VLM may not be universally advantageous, especially when textual prompts are weak. The reliance on thoroughly selected datasets introduces further limitations. MammoBLIP Schultheiss et al., [19], for example, shows impressive results in structured report generation. However, that does not ensure the correctness of the generated reports. Similarly, Mammo-CLIP [20] and Du et al. [21] rely on specific mammography data configurations limiting scalability.

Methods

1) Dataset Description

This study combines mammography data from six publicly available datasets: King Abdulaziz University Breast Cancer Mammogram Dataset (KAU-BCMD), Annotated Digital Mammograms and Associated Non-Image Datasets (ADMANI), Cohort of Screen-age Women – Case control (CSAW-CC), VinDr-Mammo, Newfoundland and Labrador Breast Screening (NLBS), and Chinese Mammography Database (CMMD). The class distribution is shown in Figure 1. Each dataset provides standard craniocaudal (CC) and mediolateral oblique (MLO) views for both breasts.

The selected cohorts enhance representativeness and diversity across multiple axes:

- Geography & Population: Saudi Arabia (KAU-BCMD), Australia (ADMANI), Sweden (CSAW-CC), Vietnam (VinDr-Mammo), Canada (NLBS), and China (CMMD).
- Scanner Vendors & Technology: IMS Giotto (KAU-BCMD), Siemens/Hologic/Philips/Fujifilm/Sectra Imtec/Konica Minolta (ADMANI), Hologic (CSAW-CC), Siemens/IMS/Planmed (VinDr-Mammo), GE Senographe Essential (NLBS), GE Senographe DS and Siemens Mammomat Inspiration (CMMD).
- Acquisition Context: Primarily screening-focused (CSAW-CC, NLBS, ADMANI) vs. diagnostic-enriched (VinDr-Mammo, CMMD, KAU-BCMD) protocols.

This heterogeneity promotes generalizability by exposing models to real-world variations in image contrast, resolution, noise, and pathology presentation across diverse ethnicities, ages, and breast densities. However, underrepresented regions (e.g., Africa, South Asia) and low-prevalence subtypes (e.g., lobular carcinoma, microcalcification-only cases) remain absent, potentially limiting performance on rare clinical scenarios or non-Asian/Caucasian populations.

For consistency, all samples were standardized into a binary classification schema:

- Normal: No radiologically detected abnormality
- Abnormal: Includes both benign and malignant findings

To address class imbalance, random undersampling was applied to the majority (normal) class. This yielded a partially balanced dataset, where approximately 45% abnormal and 55% normal cases, improving minority class representation while maintaining dataset diversity. Other balancing methods have been tested, but were discarded, since they were just duplicating abnormal cases without enriching the distribution. Random undersampling was preferred over oversampling to preserve original data diversity and avoid artificial replication of rare abnormal cases, which could inflate performance on duplicated samples while reducing generalizability across scanners and populations.

2) Preprocessing and Augmentation

Each mammogram pair (CC and MLO) corresponding to the same breast was combined into a single dual-view image by horizontal concatenation (see Figure 2).

These augmentations enhance robustness against positional, and illumination variations commonly observed in mammography.

3) Model Architectures and Adaptation Strategies

We trained and evaluated three vision-language models—Quantized Qwen2-VL-2B, Quantized SmolVLM (Idefics3-based), and MammoCLIP—alongside a convolutional neural network baseline, EfficientNet-B4.

For the general-purpose VLMs (Qwen2-VL-2B and SmolVLM), we explored two distinct adaptation strategies. The first was Full Supervised Fine-Tuning (SFT), which involved end-to-end fine-tuning of the model using LoRA adapters applied to both the vision and language modules. This approach allowed for parameter-efficient fine-tuning while leveraging the full capacity of the model. The second strategy, Linear Probing (LP), took a more resource-conscious approach by freezing the language module and the vision encoder and fine tuning only the classifier block (see Figure 3).

Table 1 Data Augmentations

Augmentation Type	Parameters / Range
Random horizontal flip	P(flip) = 0.5
Rotation	±10°
Brightness adjust	±15%
Contrast adjust	±15%

In contrast to these general-purpose models, MammoCLIP represents a domain-specific VLM pre-trained exclusively on mammography image-text pairs. For this model, we adopted the linear probing approach, maintaining the frozen vision and text encoders while fine-tuning a lightweight classifier.

Finally, as a non-VLM baseline, we employed EfficientNet-B4, a high-performing convolutional neural network architecture. This model was trained from scratch on our concatenated dual-view mammograms using the standard binary cross-entropy loss, serving as a benchmark to assess the added value of multimodal approaches compared to traditional CNN-based methods.

Table 2 Model architecture. Note that MammoCLIP was not trained using SFT, since it was pretrained on mammographic data, see third column. EffNet-B4 classifier is not complex enough for linear probing.

Model	Type	Pretrained On	Trainable # parameters on SFT	Trainable # parameters on LP	Domain-specific
Qwen2-VL	VLM	Image-Text web pairs	14M/2B	1.5M	No
SmolVLM	VLM	Image-Text web pairs	11M/1.2B	0.7M	No
MammoCLIP	VLM	Mammo-grams	-	1M	Yes
EffNet-B4	CNN	ImageNet	17M	-	No

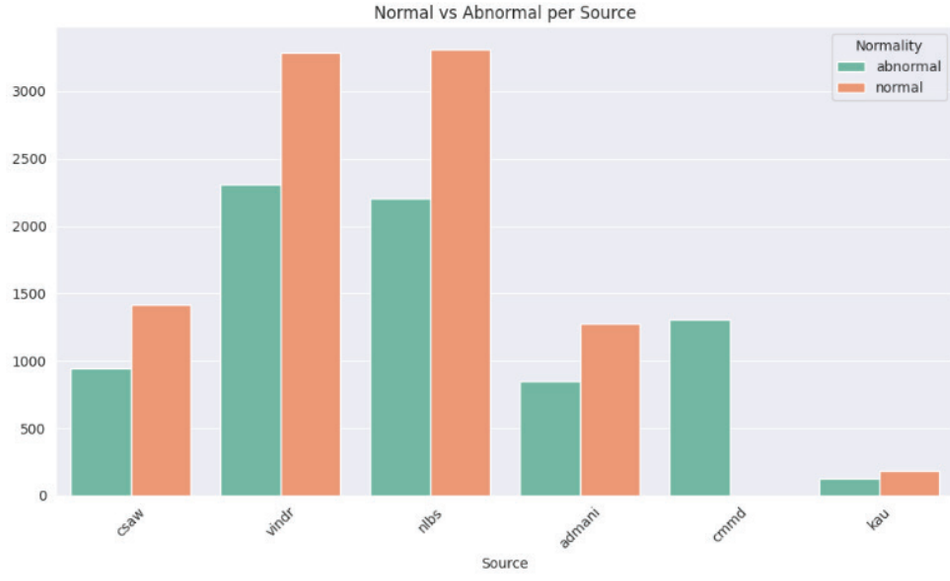


Figure 1 – Class distribution of the targets: Normal (Orange) vs Abnormal (Green)

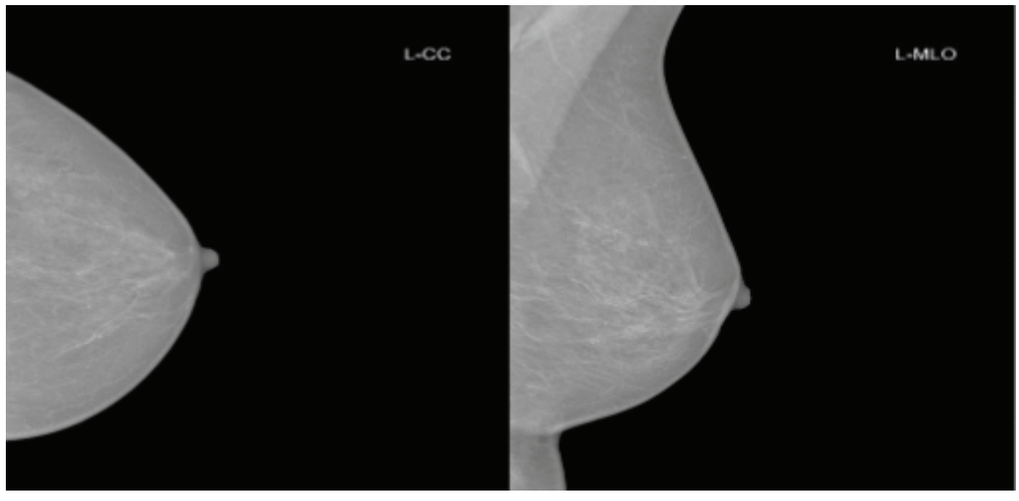


Figure 2 – Example of dual-view mammogram image (CC + MLO concatenation)

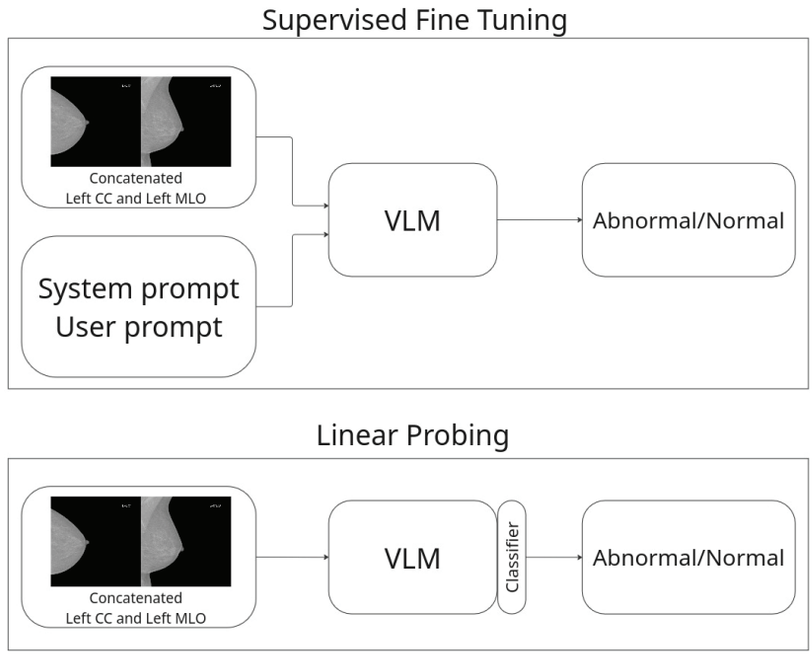


Figure 3 – Training mode pipelines

Table 3 Training hyperparameters

Hyperparameter	Value/Method
Epochs	5
Optimizer	AdamW
Learning rate	1e-4 for classification layers, 1e-5 for unfrozen vision encoder
Weight Decay	0.01
Batch size	32
Scheduler	Cosine Annealing (T_max = total steps × epochs)
Loss function	BCEWithLogitsLoss for binary classification in LP, CrossEntropyLoss for SFT experiments

4) Training Setup

The training hyperparameters are displayed in Table 3:

CrossEntropyLoss for SFT experiments

Training was conducted with mixed-precision (FP16) using PyTorch AMP for memory efficiency on a single NVIDIA RTX 4080 (16GB) GPU. The training pipeline was outlined in Figure 4.

5) Evaluation Protocol

Models were evaluated on a held-out test set using the following metrics: Accuracy (ACC), Precision, Recall (Sensitivity), F1-score, ROC-AUC.

Binary predictions were generated by applying a 0.5 threshold to the sigmoid outputs.

6) Implementation and Reproducibility

Experiments were implemented in PyTorch 2.7 with HuggingFace Transformers 4.53. Full SFT was facilitated using Unsloth, which optimizes memory usage during fine-tuning. Pretrained weights were sourced from HuggingFace (Qwen2, SmolVLM) and the official MammoCLIP repository.

Results

Table 4 summarizes the performance of all models under different adaptation strategies. Among general-purpose VLMs, Qwen2-VL-2B under full fine-tuning (FT) achieved the highest precision (97.6%) but exhibited extremely low recall (17.0%), leading to an overall F1-score of only 0.2903. This imbalance resulted in a ROC-AUC of 0.5835, indicating limited generalization capability despite its high confidence in normal cases.

SmolVLM, a smaller and more efficient VLM, achieved a more balanced performance under full fine-tuning, reaching an F1-score of 0.5101 and a ROC-AUC of 0.6304. Its Linear Probing (LP) variant delivered slightly lower F1 (0.3834) but achieved the highest ROC-AUC among VLMs (0.6368), suggesting that freezing the language module while updating the vision encoder can yield competitive results.

The domain-specific Mammo-CLIP model demonstrated the best overall performance among VLMs, achieving an F1-

Table 4 Results on the test set (normal/abnormal). FT=Fine-Tuning, LP=Linear Probing

Model	Training Mode	Accuracy	Precision	Recall	F1	ROC-AUC
Qwen2-VL-2B	FT	0.6249	0.9759	0.1705	0.2903	0.5835
Qwen2-VL-2B	LP	0.6298	0.7597	0.2590	0.3863	0.6296
SmolVLM	FT	0.6534	0.7003	0.4012	0.5101	0.6304
SmolVLM	LP	0.6287	0.7580	0.2566	0.3834	0.6368
Mammo-CLIP	LP	0.6455	0.7110	0.3572	0.4755	0.6906
EffNet-B4	FT	0.6279	0.5887	0.5735	0.5810	0.6780

score of 0.4755 and the highest ROC-AUC (0.6906) despite using only LP. This reflects the effectiveness of domain-specific pretraining in enhancing classification performance for mammography.

For comparison, the CNN baseline EfficientNet-B4, trained from scratch, achieved the highest F1-score overall (0.5810) and a ROC-AUC of 0.6780.

Discussion

These results highlight key insights into adapting vision-language models (VLMs) for mammography classification. First, general-purpose VLMs struggle to achieve balanced sensitivity and specificity on limited medical data. Despite full fine-tuning, Qwen2-VL-2B exhibited high precision (97.6%) but extremely low recall (17.0%), yielding an F1-score of only 0.2903 and ROC-AUC of 0.5835—underscoring the risk of missing positive cases in clinical screening.

Second, smaller architectures like SmolVLM showed greater adaptability. Under full fine-tuning, it achieved a balanced F1-score of 0.5101 and ROC-AUC of 0.6304. Notably, linear probing (LP) delivered competitive results (F1 = 0.3834, ROC-AUC = 0.6368), suggesting that freezing the language module while updating the vision encoder offers a cost-effective alternative for resource-constrained settings.

Domain-specific pretraining proved decisive. MammoCLIP, pretrained on mammography-aligned image-text pairs [12,20], achieved the highest VLM performance under LP (F1 = 0.4755, ROC-AUC = 0.6906) despite using only a lightweight classifier. This demonstrates that leveraging domain knowledge significantly enhances classification robustness, even without language module tuning.

In this comparative evaluation, our LP adaptation of MammoCLIP—originally pretrained on limited mammography corpora (e.g., RSNA, UPMC [12,20])—on a diverse six-dataset cohort achieved a ROC-AUC of 0.6906, demonstrating improved generalizability over its reported single-dataset zero-shot (AUC 0.61–0.68) and linear probing (AUC 0.90–0.94) configurations on VinDr [12]. This multi-source scaling enhances robustness to vendor and population shifts beyond prior single-dataset or zero-shot setups [15,16]. While generative models like MammoVLM [14] excel in radiologist-scored report quality using custom evaluation protocols, direct metric comparison is limited. Similarly, vision-only models like ConvNeXt [18] outperform VLMs in breast density assessment but lack multimodal reasoning capabilities.

For reference, our EfficientNet-B4 CNN baseline achieved the highest F1-score (0.5810) and ROC-AUC of 0.6780, confirming its strength in structured imaging tasks. However, VLMs offer unique clinical advantages beyond classification. SmolVLM, for instance, matched EfficientNet-B4's ROC-AUC while retaining language generation capability, enabling structured report drafting (Table 5), explainability, and interactive decision support—features unavailable in traditional CNNs.

These multimodal capabilities could justify deploying quantized VLMs in clinical workflows, even when raw classification accuracy is slightly lower. By scaling domain-specific models across heterogeneous public datasets via efficient probing, our work bridges predictive performance with practical AI-assisted radiology tools.

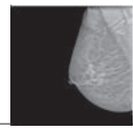
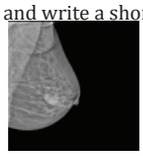
While VLMs produce structured reports, risks include hallucination (e.g., inventing calcifications) and overgeneralization from dense tissue. The abnormal example in Table 5 correctly identifies calcifications but lacks laterality and BI-RADS grading. Clinical deployment requires radiologist oversight and prompt engineering to ensure safety and specificity.

Conclusion

This study investigated the adaptation of vision-language models (VLMs) for mammography classification, comparing them to a strong CNN baseline. Our experiments demonstrate that while traditional CNNs such as EfficientNet-B4 still achieve the highest overall F1-score, VLMs provide competitive performance, particularly when leveraging domain-specific pretraining and efficient adaptation strategies.

Among VLMs, Mammo-CLIP—pretrained on mammography-related data—achieved the best overall results, with a ROC-AUC of 0.6906 under Linear Probing (LP), outperforming general-purpose VLMs even without language module fine-tuning. This underscores the value of domain-specific pretraining for high-stakes medical applications. SmolVLM, despite its smaller size, exhibited strong adaptability, delivering balanced performance under full fine-tuning and competitive

Table 5 Example of report generation using SmolVLM

Prompt	Generated report
1) "<image>\n" 2) "You are a medical AI assistant. Analyze the given mammogram image, classify it into normal vs abnormal and write a short radiology report(3 sentences) describing abnormalities, if any." 	normal. radiology report: ''No findings. Examination done. No other tests or interventions required. Follow-up as per local guideline. Risk factors to address: age, race, family history. Consultation not required. Consultation as per patient's choice of care.''
1) "<image>\n" 2) "You are a medical AI assistant. Analyze the given mammogram image, classify it into normal vs abnormal and write a short radiology report (3 sentences) describing abnormalities, if any." 	abnormal. The mammogram image is of a breast. The breast tissue is dense and has a lot of calcium deposits.

ROC-AUC scores with LP, highlighting the potential for cost-effective deployment in resource-constrained settings.

Although Qwen2-VL-2B underperformed compared to smaller or domain-specialized models, these findings suggest that scale alone does not guarantee success in medical imaging tasks. Instead, task alignment and efficient adaptation strategies are more critical.

Beyond classification performance, VLMs introduce unique capabilities—such as report generation, natural language interaction, and multimodal reasoning—that are not possible with traditional CNNs. These features can enhance clinical workflows by improving explainability, facilitating structured reporting, and enabling decision support, making VLMs an attractive direction for future research and deployment.

Author Contributions: Conceptualization, _____; methodology, _____; validation, _____; formal analysis, _____; investigation, _____; resources, _____; data curation, _____; writing – original draft preparation, _____; writing – review and editing, _____; visualization, _____; supervision, _____; project administration, _____; funding acquisition, _____. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: The authors thank two anonymous peer reviewers for their helpful suggestions.

Funding: This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan, grant number BR24993145.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

Results

1. World Health Organization. Breast cancer. Geneva: World Health Organization; 2023. Available from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer
2. Lehman CD, Arao RF, Sprague BL, Lee JM, Buist DS, Kerlikowske K, Henderson LM, Onega T, Tosteson AN, Rauscher GH, Miglioretti DL. National performance benchmarks for modern screening digital mammography: update from the Breast Cancer Surveillance Consortium. *Radiology*. 2017;283(1):49-58. https://doi.org/10.1148/radiol.2016161174
3. Liberman L, Menell JH. Breast imaging reporting and data system (BI-RADS). *Radiol Clin North Am.* 2002;40(3):409-430. https://doi.org/10.1016/s0033-8389(01)00017-3
4. Abdikenov B, Zhaksylyk T, Shortanbaiuly O, Orazayev Y, Makhanov N, Karibekov T. Future of breast cancer diagnosis: A review of DL and ML applications and emerging trends for multimodal data. IEEE Access. 2025;13:1-15. https://doi.org/10.1109/ACCESS.2025.3585377
5. Li X, Li L, Jiang Y, Wang H, Qiao X, Feng T, Luo H, Zhao Y. Vision-language models in medical image analysis: From simple fusion to general large models. *Inf Fusion.* 2025;118:102995. https://doi.org/10.1016/j.inffus.2025.102995
6. Qin Z, Yi H, Lao Q, Li K. Medical image understanding with pretrained vision–language models: A comprehensive study. In: Proceedings of the Eleventh International Conference on Learning Representations (ICLR); 2023. Available from: https://openreview.net/pdf?id=txlWziuCE5W
7. Zhang J, Huang J, Jin S, Lu S. Vision-Language Models for Vision Tasks: A Survey. *IEEE Trans Pattern Anal Mach Intell.* 2024;46(8):5625-5644. https://doi.org/10.1109/TPAMI.2024.3369699
8. Güneş YC, Cesur T, Çamur E, Karabekmez LG. Evaluating text and visual diagnostic capabilities of large language models on questions related to the Breast Imaging Reporting and Data System Atlas 5th edition. *Diagn Interv Radiol*. 2025;31(2):111-129. https://doi.org/10.4274/dir.2024.242876
9. Haver HL, Bahl M, Doo FX, Kamel PI, Parekh VS, Jeudy J, Yi PH. Evaluation of multimodal ChatGPT (GPT-4V) in describing mammography image features. *Can Assoc Radiol J.* 2024;75(4):947-949. https://doi.org/10.1177/08465371241247043
10. Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J, Krueger G, Sutskever I. Learning transferable visual models from natural language supervision. In: Proceedings of the 38th International Conference on Machine Learning. PMLR; 2021. p. 8748-8763. Available from: https://proceedings.mlr.press/v139/radford21a.html
11. Wang Z, Wu Z, Agarwal D, Sun J. MedCLIP: Contrastive learning from unpaired medical images and text. In: Proceedings of the *Conference on Empirical Methods in Natural Language Processing (EMNLP);* 2022. p. 3876–3887. https://doi.org/10.18653/v1/2022.emnlp-main.256
12. Ghosh S, Poynton CB, Visweswaran S, Batmanghelich K. Mammo-CLIP: A vision–language foundation model to enhance data efficiency and robustness in mammography. In: Linguraru MG, et al., editors. Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. *Lecture Notes in Computer Science. Springer;* 2024. p. 623-632. https://doi.org/10.1007/978-3-031-72390-2_59
13. Yan Z, Zhang W, Yang Y, Saha A. Multi-distribution mammogram classification: Leveraging CLIP for enhanced generalization across diverse datasets. In: Proceedings of the 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). *IEEE*; 2025. p. 1-5. https://doi.org/10.1109/ISBI60581.2025.10981260
14. Cao Z, Deng Z, Ma J, Hu J, Ma L. MammoVLM: A generative large vision–language model for mammography-related diagnostic assistance. *Inf Fusion.* 2025;102998. https://doi.org/10.1016/j.inffus.2025.102998
15. Yan Z. Enhancing zero-shot learning with CLIP for multi-distribution mammogram classification [Master's thesis]. Western University; 2025. Available from: https://ir.lib.uwo.ca/etd/10816
16. Vo HQ, Wang L, Wong KK, Ezeana CF, Yu X, Yang W, Chang J, Nguyen HV, Wong STC. Frozen large-scale pretrained vision–language models are the effective foundational backbone for multimodal breast cancer prediction. IEEE J Biomed Health Inform. 2025;29(5):3234-3246. https://doi.org/10.1109/JBHI.2024.3507638

17. Guo X, Chai W, Li SY, Wang G. LLaVA-Ultra: Large Chinese Language and Vision Assistant for Ultrasound. In: Proceedings of the 32nd ACM International Conference on Multimedia. ACM; 2024. p. 8845-8854. https://doi.org/10.1145/3664647.3681584
18. Molina-Román Y, Gómez-Ortiz D, Menasalvas-Ruiz E, Tamez-Peña JG, Santos-Díaz A. Comparison of ConvNeXt and vision–language models for breast density assessment in screening mammography. In: Proceedings of the 2025 IEEE 38th International Symposium on Computer-Based Medical Systems (CBMS). *IEEE*; 2025. p. 636–641. https://doi.org/10.1109/CBMS65348.2025.00132
19. Schultheiss T, Walther L, Maunz J, Weninger L, Maier A, Wirkert S. MammoBLIP: End-to-end mammography report generation using a curated, standardized, multi-institutional public dataset [Manuscript under review]. Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg; 2024. Available from: https://lme.tf.fau.de/wp-content/uploads/2020/08/MammoBLIP__End_to_End_Mammography_Report_Generation_Using_a_Curated__Standardized_Multi_Institutional_Public_Dataset-1.pdf
20. Chen X, Li Y, Hu M, Salari E, Chen X, Qiu RLJ, Zheng B, Yang X. Mammo-CLIP: Leveraging contrastive language-image pre-training (CLIP) for enhanced breast cancer diagnosis with multi-view mammography. arXiv preprint. 2024;arXiv:2404.15946. https://doi.org/10.48550/arXiv.2404.15946
21. Du Y, Onofrey J, Dvornek NC. Multi-view and multi-scale alignment for contrastive language-image pre-training in mammography. arXiv preprint. 2024;arXiv:2409.18119. https://doi.org/10.48550/arXiv.2409.18119

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17473

Surgery and Survival in Patients with Intrahepatic Cholangiocarcinoma: A Bayesian Inference

Ruslan Akhmedullin[1], Zhandos Burkitbayev[2], Tair Koishibayev[2], Zhanat Spatayev[2], Abylaikhan Sharmenov[2], Oxana Shatkovskaya[2], Dinara Zharlyganova[2], Almira Manatova[2], Zhuldyz Kuanysh[2], Sanzhar Shalekenov[2]

[1]Department of Medicine, Nazarbayev University School of Medicine, Astana, Kazakhstan
[2]Department of Science, National Research Oncology Centre, Astana, Kazakhstan

Received: 2025-10-16.
Accepted: 2025-11-27.

J Clin Med Kaz 2025; 22(6): 96–102

Corresponding author:
Zhandos Burkitbayev
Email: 6541626@gmail.com.
ORCID: 0009-0000-4859-1637.

ABSTRACT

Background: Conventional approaches often struggle to draw inferences from limited data. We aimed to determine whether Bayesian perspectives improve computation and increase clarity when inferring from small series.

Methods: We conducted a retrospective longitudinal study. The participants were who underwent curative liver resection for ICC. Bayesian proportional hazards models were used to analyse postsurgical survival. Three scenarios were used to model overall survival; all estimates were then compared with the largest meta-analysis.

Results: The final cohort included 14 patients with resectable ICC with median follow-up duration of 14.68 months (IQR 10.77 - 19.84). While the model with non-informative priors revealed unstable estimates, the models with informative and less informative priors demonstrated consistent and reliable effects. The estimates from the latter models aligned with those reported in the meta-analysis. The posterior probabilities of null value exceedance (i.e., HR > 1.0) for age, sex, lymph node metastasis, and positive surgical margin were 98%, 93%, 100%, and 48%, respectively, albeit (in frequentist terms) non-significant for all but lymph node metastasis.

Conclusion: The Bayesian framework offers a promising strategy that overcomes the limitations of the frequentist approach, contributing to the advancement of inferences from limited data.

Keywords: ICC; Survival; Mortality; Bayesian Analysis; Methodology;

Introduction

Liver cancer remains one of the leading causes of cancer-related deaths worldwide, with notably high estimates in developing regions [1]. It represents a grouping disorder, with cholangiocarcinoma being the second most common primary malignant tumour [1, 2]. Cholangiocarcinoma is broadly classified as intrahepatic cholangiocarcinoma (ICC) or extrahepatic cholangiocarcinoma (ECC) depending on the anatomical location [2, 3]. The former is of particular interest because of increasing mortality patterns worldwide [3].

Most ICC cases are diagnosed at advanced stages, where a dismal prognosis is particularly evident [4, 5]. Complete surgical resection is often the only potential rescue for prolonged survival in these patients [6, 7]. Despite the tumour resection, however, the long-term survival remains low, primarily due to tumour recurrence and lymph node metastases [4, 8, 9]. Admittedly, survival estimates are heterogeneous: existing evidence [10] suggests that 3-year and 5-year survival ranges from 16% to 65%, and 5% to 56%, respectively.

The existing findings on surgical outcomes in patients with ICC primarily come from retrospective studies, which lack methodological rigor [10, 11]. What is more, estimations in the living meta-analysis [10] were restricted to studies from countries with large populations, further setting aside evidence from small series (e.g., n <20). However, in the subgroup analysis of the impact of age on the survival, a study with the lowest sample size (n=22) revealed non-statistically significant effect of age (HR 1.30, 95% CI: 0.92-1.84). Whereas the one with the largest sample size (n=733) saw significant association (HR 1.10, 95% CI: 1.01-1.20). However, the (log) difference between these studies was 0.17, with a standard error of 0.18. Hence, despite the discrepancies, the difference between their estimates is not itself statistically significant. This questions the relevance of the selection based on the sample size, as these studies provide complementary evidence, and make a small data valuable in its own right. On the other hand, it reflects the asymptomatic nature of ICC and its overall low incidence, where locations with a low population at risk face low incident cases. The ignorance of small studies may contribute to the limited generalizability and risk of overlooking meaningful insights into surgical outcomes. However, a notable difficulty arises when analysing such data, as they may create more noise than signals, potentially leading to overconfident conclusions. These highlight the importance of small series and the necessity of appropriate statistical models to draw inferences from them.

The Bayesian approach allows to integrate prior findings to build upon existing knowledge rather than modelling the data anew [12]. This feature often makes Bayesian framework preferable when working with small samples [13], with growing evidence on its advantages for the limited data [14, 15]. Previous studies have employed uniform analytical approach and disregarded the priors, pretending that their research is unique (unless it is a meta-analysis) [16]. We instead suggest transitioning from that "one-size-fits-all" approach, building inferences acknowledging the empirical evidence.

This study aimed to analyse the overall survival (OS) in a small cohort of patients with ICC who underwent curative liver resection and to investigate the impact of tumour-specific factors on OS. We employ a Bayesian framework and illustrate how the prior choice affects the models' coefficients. We begin by setting a model using non-informative priors, followed by a more complex model incorporating strong priors, and contrast them with the meta-analysis estimates [10]. We then discuss the findings and their future implications.

Methods

Study design and case definitions

A retrospective longitudinal study was conducted to evaluate the OS in the ICC cohort following the surgical treatment. We included all the histologically confirmed cases who have had a liver resection at National Research Oncology Centre in the Republic of Kazakhstan (NROC) [17]. The TNM stage grouping was computed in accordance to the criteria of the eighth edition of the AJCC Cancer Staging Manual [18].

Data source and population

This study departed data from the NROC between January 2022 and July 2025. Every patient had a standard clinical examination, an assessment of serum laboratory tests, computed tomography and/or magnetic resonance cholangiopancreatography as part of their pre-operative workup. Patients could only have surgery if the tumour could be removed entirely while leaving a functional liver remnant that had enough hepatic venous outflow and vascular inflow. Data collection and analysis were performed during June – July 2025.

Follow-up and outcome assessment

In compliance with the local guidelines, patients were monitored every three months. The censoring date was defined as the last date the patient was contacted by a clinician, or as the date of death. OS time was computed as the difference (in months) between the date of surgery, and death, last date that a patient known to be alive or the follow-up cut-off date (July 18, 2025), whichever came first.

Rationale for using Bayesian analysis

Our primary focus was on the statistical inferences of the association between patient and tumour characteristics and survival, magnitude of the effect, and its uncertainties, rather than significance testing. We were also interested in the extent to which strong priors (external evidence) affect the estimates, which is challenging in the conventional frequentist approach. In addition, the existing literature suggests that the estimates were fluctuating when exploring the effects of clinical factors on survival [10]. This may be an artifact of the sample size, which was mostly small. However, when priors are informative and data are weak (i.e., limited, noisy, restricted), Bayesian perspective can bring a distinct impact on the posterior estimates, further improving inferences, their interpretation, and the precision of effect sizes (e.g., sign, magnitude) [19-21].

Variable selection and prior identification

Previous studies have often implemented backward elimination [22, 23] and stepwise regression [8, 9] for variable selection. However, such variable filtering may result in unstable inferences that require substantially more cases to ensure adequate events per variable ratio and introduce systematic errors [24, 25], which is even more challenging given ICC remains uncommon. Therefore, we did not pre-filter the variables or eliminate weak effects. Instead, we relied on expert background knowledge, including clinically important variables, and incorporated external evidence of their effects [10]. This approach is known to reduce bias due to systematic errors (unmeasured confounding) [26].

To analyse the survival factors, we used the following variables: age (continuous scale), sex (male vs. female), positive surgical margin (R1 vs. R0), and lymph node metastases (N1 vs. N0). Continuous variables (e.g., age and follow-up) are reported as medians and interquartile ranges (IQR), while categorical variables are reported as proportions. The distribution of the priors for the tumour-specific covariates was extracted from a recent meta-analysis [10]. We treated its estimates as a true effect size, since the study used strict methodological approaches to evaluate and report survival estimates in patients with ICC following curative surgery (852 citations by August, 2025). We used a log-transformed estimate, and its scale (standard deviation, SD) was approximated using the range of the 95% CI, and used the following formula: [upper limit – lower limit / 3.92] [27]. Given the limited number of events per variable and the associated risk of unstable estimates in multivariable analysis, we report only estimates from the univariate analyses.

Statistical analysis

We employed Bayesian proportional hazards regression models, which was motivated by the following reasons: the measure (Hazard Ratio, HR) offers an intuitive interpretation

of coefficients and consistently used across the studies. It also ensures simplicity in incorporating prior information into the models without any conversions (e.g., from log-odds to log-hazards). The 95% credible intervals (CrI) were constructed directly from the 2.5th and 97.5th percentiles of the posterior distribution of the corresponding models. Our analysis used weakly (default) informative (Model 1) and informative (Model 2) prior distributions to model the data. All the models retrieved four chains of 2,000 Markov Chain Monte Carlo samples, the first 1,000 samples of which were discarded as a warm-up. Model convergence was evaluated using the R-hat diagnostic and effective sample size statistics. Data cleaning, data management, and statistical analysis were performed using R-software (version 4.5.0). All methods were performed in accordance with the STROBE Guidelines [28].

Sensitivity analysis.

We performed a post-hoc analysis (Model 3) to test the reliability of our results for different prior choices. Initial prior values were extracted from the meta-analysis [10], in which most reported effects were moderately positive. Recognizing their potential impact on the posterior estimates, we decreased their precision (effect size) while keeping the scale (standard error) same to see how the estimates changed given the less informative priors.

Results

Overall, this study included data from 14 patients with ICC whom undergone a curative liver resection. All the patients had a lymphadenectomy. There were nine male (64.29%) and five female (35.71%) patients. The median age was 62 years (IQR 59 - 66) and median follow-up duration made-up 14.68 months (IQR 10.77 - 19.84). All the other clinical and demographical information is available in the Supplemental Table 1.

Regression analyses revealed consistent and stable estimates only for age, whereas other covariates showed fluctuations, as measured by the HR (Figure 1). Moreover, lymph node metastasis and positive surgical margins saw modelling issues, which were evident when closely examining the estimates. These were primarily driven by the limited number of events and imbalances in these predictors. However, this was only true for the model with non-informative priors (Model 1). Hence, the further inclusion of informative priors (Model 2) resolved these issues. What is more, estimates from the model with strong priors (Model 2) approached those reported in the meta-analysis (Figure 1, Supplemental Table 2).

A sensitivity analysis was conducted with priors, whose precision was decreased. Interestingly, the model with less informative priors (Model 3) revealed estimates that were largely comparable to those from Model 2, with both the sign and magnitude of the effects being almost identical. The estimates in Model 3 tended to be more stable than those in Model 1 and approached the meta-analysis (Figure 1 and Supplemental Table 2).

Finally, we evaluated the probability of null value exceedance (i.e., HR > 1.0) for each covariate. In Model 2, the posterior distributions of HR for age, sex, lymph node metastasis, and positive surgical margin were 98.1%, 93.1%, 100%, and 48.2%, respectively (Figure 2). Similarly, Model 3 showed probabilities of 78.7%, 78.1%, 100%, and 43.9%, respectively (Supplemental Table 3 and Supplemental Figure 1), whereas the probabilities were substantially lower in Model 1 (Supplemental Table 3 and Supplemental Figure 2).

Discussion

We evaluated survival in a small ICC cohort using a Bayesian approach. We found that the effects of most covariates aligned with the meta-analysis as long as the priors were strong.

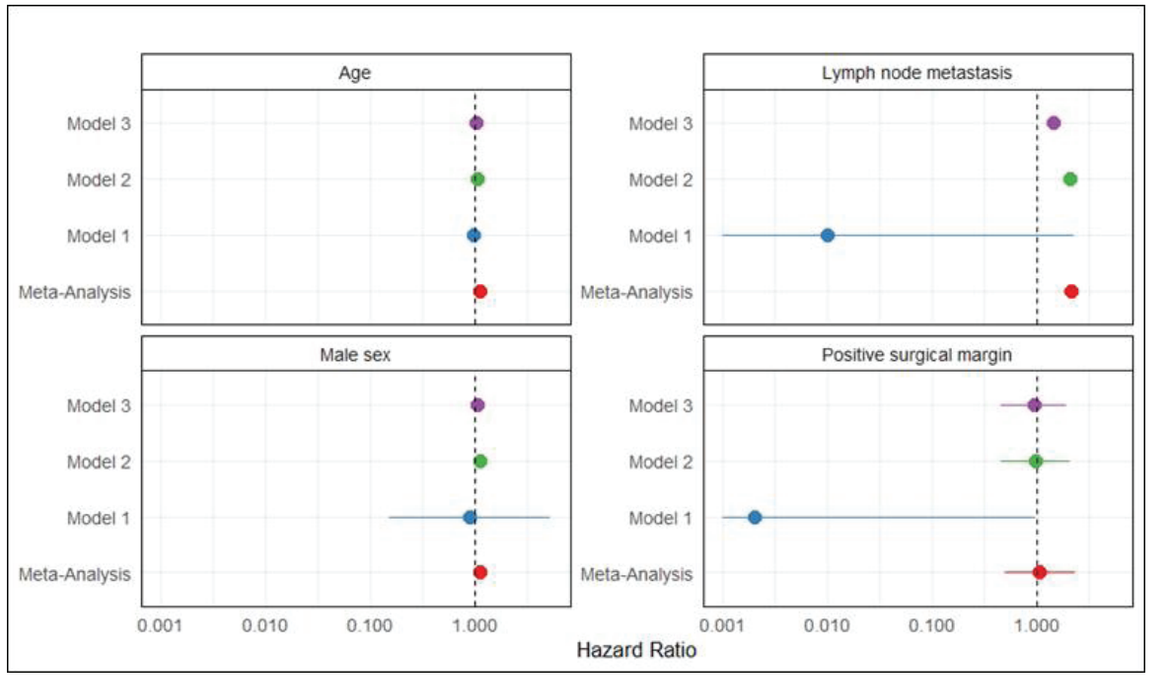


Figure 1 – Predictors of Overall Survival

CAPTION: Model 1 represents weakly informative priors; Model 2 includes strong priors; Model 3 uses informative priors, albeit decreasing the precision; the meta-analysis represents the estimates from the largest study evaluating the survival function in patients with intrahepatic cholangiocarcinoma. The estimates from the latter were treated as the true effect size.

Although the effects remained stable despite a decrease in the precision of the prior distributions, the direction and magnitude of the effects were highly sensitive to the priors. By applying Bayesian analysis, we contributed to the growing evidence demonstrating its superiority over the frequentist method, specifically for the limited data.

The frequentist paradigm has emerged as the dominant approach to statistical thinking. Previous studies [8, 22, 23, 29-31] were mostly unable to certainly report whether the covariates affected survival, with claims being primarily driven by significance testing. A large amount of literature comes from underpowered studies, which is inevitable for uncommon diseases such as ICC. What is more, at most 30% of ICC population are resectable [32]. Hence, when the sample size is small, p-values are highly sensitive to small deviations in the data, leading to ambiguous conclusions [33]. Therefore, statistical significance rarely touches on the subject of interest in such instances. Instead, we suggest taking advantages of Bayesian language to increase the reliability of inference from small data [34,35].

Bayesian framework estimates how likely different values of a parameter are by combining prior knowledge with data. Unlike frequentist approach, it provides an estimate that sampled from the posterior distribution, which represents a compromise between the priors and likelihood (data). Hence, when priors are close to the data, the posterior estimates favour the data. Similarly, the posterior distribution tends to move the location toward a prior distribution as uncertainty in the data increases (i.e., limited data, noise). In other words, when priors are strong and data are weak, Bayesian approach can have a pronounced impact on the model estimates. However, as the sample size becomes large, the data take more control over the compromise, diminishing the impact of the priors. In our study, all priors, irrespective of their precision, heavily affected the estimates. Admittedly, our posterior estimates were sensitive to priors because the information in the data dwarfed by the priors, implying the weakness of the likelihood.

The effect sizes in the meta-analysis were calculated, given a certain degree of uncertainty from an individual study [10]. In the subgroup analyses, the authors were unable to demonstrate a significant difference in OS for sex and positive surgical margin [10], which was also largely true for individual studies [8, 22, 23, 29, 30]. However, such a claim is no longer sufficient. Moreover, any scientific claim should include a probability that expresses confidence in the effect [36]. As such, we report a high (posterior) probability for age (98.1%), sex (93.1%), and lymph node metastasis (100.0%) to elevate the risks of death, albeit being non-statistically significant (in frequentist terms). However, the effect of a positive surgical margin was largely uncertain, revealing only a 48.2% probability of increased risk (i.e., HR > 1.0). These were true for both Models 2 and 3. Thus, unlike the conventional approach, Bayesian analysis offers more computational flexibility, providing more clinically relevant information and easing interpretation.

Although our sample size may seem insufficient, the cohort is the largest of its kind for patients with ICC in our region. Additionally, Bayesian estimation provides reliable results even when sample sizes are small [14, 15], with informative priors being exceptionally useful in such instances. As such, we revealed unstable estimates with weakly informative priors. Further, inclusion of strong priors improved the computation, providing both stable and plausible estimates. This was also supported by the sensitivity analysis, where the HR saw just a slight fluctuation.

We believe that our findings remain reliable and useful regardless of the limited follow-up period. The Cox regression has been consistently applied across the previous studies due to its convenience in analysing survival function. It requires hazards to be proportional (constant) for groups that are under comparison. However, it is rarely the case in practice, especially when the treatment/exposure effect (naturally) changes over time [37]. Hence, the estimate (HR) reflects a weighted average effect, rather than a constant effect. As such, incorporating the external evidence to our model, we controlled the fluctuations

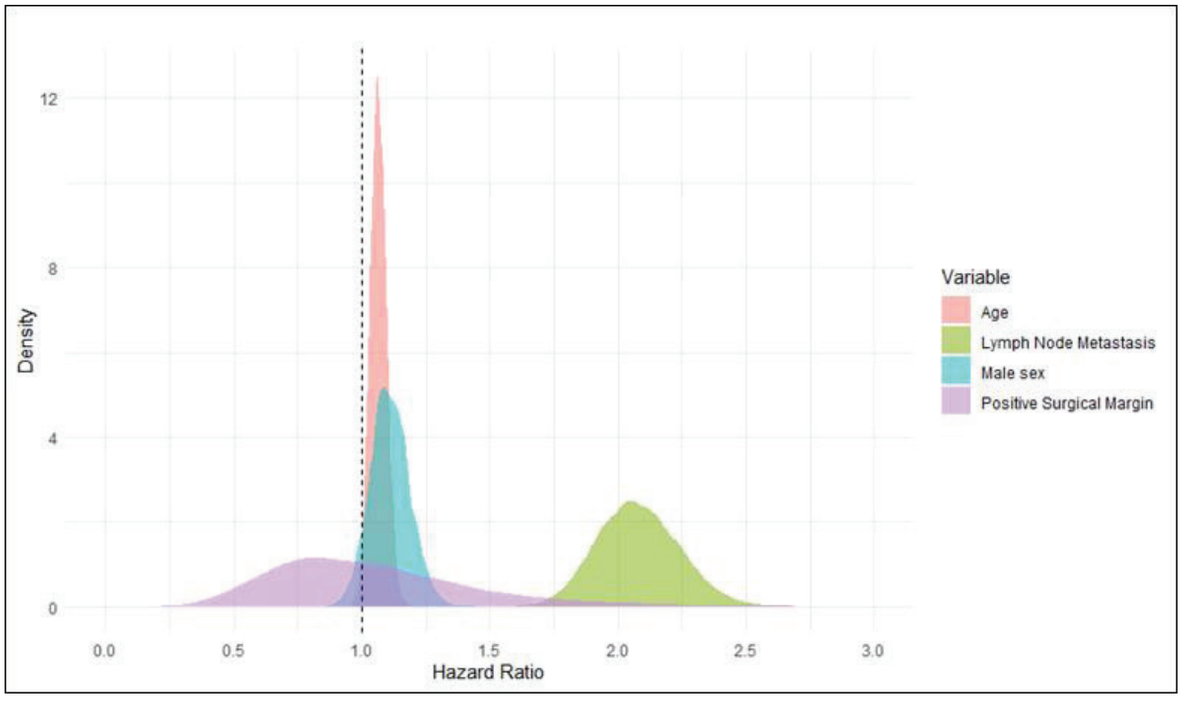


Figure 2 – Posterior probability of exceeding the Hazard Ratio of 1 for Model 2

over time and revealed reliable effect size, which is evident when contrasting estimates to the meta-analysis [10].

Finally, in the largest meta-analysis [10], 13% of studies were excluded due to low sample size (<20). However, the ignorance of these studies might lead to overconfident estimates (i.e., underestimating errors). This is particularly evident when looking at the funnel plot more closely, which revealed figures clustering toward the top and asymmetric distributions for most of the outcomes. Hence, we emphasize that small series are valuable in their own right and encourage the application of more customized statistical strategies to communicate uncertainty. Bayesian modelling is beneficial, allowing for the stabilization of estimates even for noisy data. Furthermore, rather than testing if an underlying effect is non-zero, one may access more hidden, clinically important information. In addition, the Bayesian approach allows the incorporation of external evidence, setting a range of plausible values into statistical models, and computing the probability of a meaningful effect [38]. These alleviate subsequent contrasting findings and check if an estimate is under-or-over-estimated and its consistency with the existing evidence (e.g., meta-analysis and RCT). Taken together, we believe that the Bayesian language can and should also be proposed as the default approach for limited data, or at least as a required approach for post-hoc estimations.

Limitations

This study has several limitations. Admittedly, our suggestions were based on limited data, with a handful of predictors being modelled. Efforts to contact both the authors of previous work and other hospitals in the country to acquire additional information were unsuccessful. Therefore, we may have had model specification issues, which means that our estimates were ignorant of other clinically important variables. Additionally, given the decay of priors in larger samples, it would be interesting to assess whether bringing more data proves or disproves our findings. Finally, informative priors for covariates other than those used can be included in multivariable analysis to test the stability of the estimates.

Conclusions

The sample size in studies on rare diseases is fairly small. Research evaluating outcomes using such data faces analytical challenges. We suggest that the Bayesian approach is a promising strategy that may contribute to advancing inferences from the limited data. We present a pioneering attempt to model ICC outcomes using a fully Bayesian approach and found that it overcomes the limitations of conventional approaches in different ways. We believe that Bayesian analysis can be proposed as a default approach when working with small data or modelling outcomes for rare diseases.

Abbreviations

ICC – Intrahepatic Cholangiocarcinoma;
RCT – Randomized Controlled Trial;
HR – Hazard Rate;
IQR – Interquartile Range;
NROC – National Research Oncology Centre;
STROBE – Strengthening the Reporting of Observational Studies in Epidemiology;
ECC – Extrahepatic Cholangiocarcinoma;
CI – Confidence Interval;
CrI – Credible Interval;
OS – Overall Survival.

Supplementary materials

The Supplementary information includes tables:

• Supplementary Table 1. Baseline characteristics of the study cohort;

Supplementary Table 2. Association between the covariates and survival in the study cohort;

Supplementary Table 3. Posterior probability of exceeding the hazard ratio of 1;

as well as figures:

• Supplementary Figure 1. Posterior probability of exceeding the hazard ratio of 1 for Model 1 (weakly informative priors);

Supplementary Figure 2. Posterior probability of exceeding the hazard ratio of 1 for Model 3 (sensitivity analysis).

This supplemental material has been provided by the authors to give readers additional information about their work.

The file can be accessed using:https://www.editorialpark.com/download/article-supp/733/Supplement-1.docx.

Author Contributions: Conceptualization, R.A. and Z.B.; methodology, O.S., Z.K., A.M.; formal analysis and investigation, D.Z., Z.K., Z.S.; writing – original draft preparation, S.S., T.K., Z.S., A.S.; writing – review and editing, Z.B., S.S., D.Z., Z.B.; supervision, Z.B. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

Ethics approval consent to participate: This study was approved by the Institutional Review Ethics Committee of the NROC (NROC-IREC – 034, July 30, 2025). Given the retrospective nature of the study and the fact that complete anonymization of all patient data was ensured, the Institutional Review Ethics Committee waived the requirement for informed consent.

References

1. Arnold M, Abnet CC, Neale RE, Vignat J, Giovannucci EL, McGlynn KA, Bray F. Global Burden of 5 Major Types of Gastrointestinal Cancer. Gastroenterology. 2020 Jul;159(1):335-49.e15. PubMed PMID: 32247694. PMCID: PMC8630546. Epub 20200402. https://doi.org/10.1053/j.gastro.2020.02.068
2. Jiang K, Al-Diffhala S, Centeno BA. Primary Liver Cancers-Part 1: Histopathology, Differential Diagnoses, and Risk Stratification. Cancer Control. 2018 Jan-Mar;25(1):1073274817744625. PubMed PMID: 29350068. PMCID: PMC5933592. https://doi.org/10.1177/1073274817744625
3. Bertuccio P, Malvezzi M, Carioli G, Hashim D, Boffetta P, El-Serag HB, La Vecchia C, Negri E. Global trends in mortality from intrahepatic and extrahepatic cholangiocarcinoma. *J Hepatol*. 2019 Jul;71(1):104-14. PubMed PMID: 30910538. Epub 20190323. https://doi.org/10.1016/j.jhep.2019.03.013
4. Esmail A, Badheeb M, Alnahar B, Almiqlash B, Sakr Y, Khasawneh B, Al-Najjar E, Al-Rawi H, Abudayyeh A, Rayyan Y, Abdelrahim M. Cholangiocarcinoma: The Current Status of Surgical Options including Liver Transplantation. *Cancers (Basel)*. 2024 May 21;16(11). PubMed PMID: 38893067. PMCID: PMC11171350. Epub 20240521. https://doi org/10.3390/cancers16111946
5. Izquierdo-Sanchez L, Lamarca A, La Casta A, Buettner S, Utpatel K, Klümpen HJ, Adeva J, Vogel A, Lleo A, Fabris L, Ponz-Sarvise M, Brustia R, Cardinale V, Braconi C, Vidili G, Jamieson NB, Macias RI, Jonas JP, Marzioni M, Hołówko W, Folseraas T, Kupčinskas J, Sparchez Z, Krawczyk M, Krupa Ł, Scripcariu V, Grazi GL, Landa-Magdalena A, Ijzermans JN, Evert K, Erdmann JI, López-López F, Saborowski A, Scheiter A, Santos-Laso A, Carpino G, Andersen JB, Marin JJ, Alvaro D, Bujanda L, Forner A, Valle JW, Koerkamp BG, Banales JM. Cholangiocarcinoma landscape in Europe: Diagnostic, prognostic and therapeutic insights from the ENSCCA Registry. J Hepatol. 2022 May;76(5):1109-21. PubMed PMID: 35167909. Epub 20220212. https://doi org/10.1016/j.jhep.2021.12.010
6. de Jong MC, Nathan H, Sotiropoulos GC, Paul A, Alexandrescu S, Marques H, Pulitano C, Barroso E, Clary BM, Aldrighetti L, Ferrone CR, Zhu AX, Bauer TW, Walters DM, Gamblin TC, Nguyen KT, Turley R, Popescu I, Hubert C, Meyer S, Schulick RD, Choti MA, Gigot JF, Mentha G, Pawlik TM. Intrahepatic cholangiocarcinoma: an international multi-institutional analysis of prognostic factors and lymph node assessment. *J Clin Oncol*. 2011 Aug 10;29(23):3140-5. PubMed PMID: 21730269. Epub 20110705. https://doi.org/10.1200/JCO.2011.35.6519
7. Rushbrook SM, Kendall TJ, Zen Y, Albazaz R, Manoharan P, Pereira SP, et al. British Society of Gastroenterology guidelines for the diagnosis and management of cholangiocarcinoma. *Gut*. 2023 Dec 7;73(1):16-46. PubMed PMID: 37770126. PMCID: PMC10715509. Epub 20231207. https://doi.org/ 10.1200/JCO.2011.35.6519
8. Ribero D, Pinna AD, Guglielmi A, Ponti A, Nuzzo G, Giulini SM, Aldrighetti L, Calise F, Gerunda GE, Tomatis M, Amisano M, Berloco P, Torzilli G, Capussotti L. Surgical Approach for Long-term Survival of Patients With Intrahepatic Cholangiocarcinoma: A Multi-institutional Analysis of 434 Patients. *Arch Surg*. 2012 Dec;147(12):1107-13. PubMed PMID: 22910846. https://doi.org/10.1001/archsurg.2012.1962
9. Zhang XF, Beal EW, Bagante F, Chakedis J, Weiss M, Popescu I, Marques HP, Aldrighetti L, Maithel SK, Pulitano C, Bauer TW, Shen F, Poultsides GA, Soubrane O, Martel G, Koerkamp BG, Itaru E, Pawlik TM. Early versus late recurrence of intrahepatic cholangiocarcinoma after resection with curative intent. *Br J Surg*. 2018 Jun;105(7):848-56. PubMed PMID: 29193010. Epub 20171128. https://doi.org/10.1002/bjs.10676
10. Mavros MN, Economopoulos KP, Alexiou VG, Pawlik TM. Treatment and Prognosis for Patients With Intrahepatic Cholangiocarcinoma: Systematic Review and Meta-analysis. JAMA Surg. 2014 Jun;149(6):565-74. PubMed PMID: 24718873. https://doi.org/10.1001/jamasurg.2013.5137
11. Song Y, Cai M, Li Y, Liu S. The focus clinical research in intrahepatic cholangiocarcinoma. *Eur J Med Res*. 2022 Jul 11;27(1):116. PubMed PMID: 35820926. PMCID: PMC9277934. Epub 20220711. https://doi.org/10.1186/s40001-022-00791-z
12. Lammers D, Richman J, Holcomb JB, Jansen JO. Use of Bayesian Statistics to Reanalyze Data From the Pragmatic Randomized Optimal Platelet and Plasma Ratios Trial. JAMA Netw Open. 2023 Feb 1;6(2):e230421. PubMed PMID: 36811858. PMCID: PMC9947730. Epub 20230201. https://doi.org/ 10.1001/jamanetworkopen
13. McNeish D. On Using Bayesian Methods to Address Small Sample Problems. Structural Equation Modeling: A Multidisciplinary Journal. 2016 2016/09/02;23(5):750-73. https://doi.org 10.1080/10705511.2016.1186549
14. van de Schoot R, Broere JJ, Perryck KH, Zondervan-Zwijnenburg M, van Loey NE. Analyzing small data sets using Bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. Eur J Psychotraumatol. 2015;6:25216. PubMed PMID: 25765534. PMCID: PMC4357639. Epub 20150311. https://doi.org/10.3402/ejpt.v6.25216
15. Larson C, Kaplan D, Girolamo T, Kover ST, Eigsti IM. A Bayesian statistics tutorial for clinical research: Prior distributions and meaningful results for small clinical samples. *J Clin Psychol*. 2023 Nov;79(11):2602-24. PubMed PMID: 37477577. PMCID: PMC10591806. Epub 20230721. https://doi.org/ 10.1002/jclp.23570
16. Zwet Ev, Gelman A. A Proposal for Informative Default Priors Scaled by the Standard Error of Estimates. The American Statistician. 2022 2022/01/02;76(1):1-9. https://doi.org/10.1080/00031305.2021.1938225
17. Beyembetova A, Ablayeva A, Akhmedullin R, Abdukhakimova D, Biniyazova A, Gaipov A. National Electronic Oncology Registry in Kazakhstan: Patient's Journey. *Epidemiology and Health Data Insights*. 2025;1(1):ehdi004. https://doi.org/10.63946/ehdi/16385
18. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, Meyer L, Gress DM, Byrd DR, Winchester DP. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. 2017 Mar;67(2):93-9. PubMed PMID: 28094848. Epub 20170117. https://doi.org/10.3322/caac.21388
19. Gelman A, Carlin J. Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspect Psychol Sci*. 2014 Nov;9(6):641-51. PubMed PMID: 26186114. https://doi.org/10.1177/1745691614551642
20. Gustafson P, McCandless LC, Levy AR, Richardson S. Simplified Bayesian sensitivity analysis for mismeasured and unobserved confounders. Biometrics. 2010 Dec;66(4):1129-37. PubMed PMID: 20070294. https://doi.org/10.1111/j.1541-0420.2009.01377.x
21. Scharfstein DO, Daniels MJ, Robins JM. Incorporating prior beliefs about selection bias into the analysis of randomized trials with missing outcomes. Biostatistics. 2003 Oct;4(4):495-512. PubMed PMID: 14557107. PMCID: PMC2748253. https://doi.org/10.1093/biostatistics/4.4.495

22. Wang Y, Li J, Xia Y, Gong R, Wang K, Yan Z, Wan X, Liu G, Wu D, Shi L, Lau W, Wu M, Shen F. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. *J Clin Oncol.* 2013 Mar 20;31(9):1188-95. PubMed PMID: 23358969. Epub 20130128. https://doi.org/10.1200/JCO.2012.41.5984
23. Clark CJ, Wood-Wentz CM, Reid-Lombardo KM, Kendrick ML, Huebner M, Que FG. Lymphadenectomy in the staging and treatment of intrahepatic cholangiocarcinoma: a population-based study using the National Cancer Institute SEER database. HPB (Oxford). 2011 Sep;13(9):612-20. PubMed PMID: 21843261. PMCID: PMC3183445. https://doi.org/10.1111/j.1477-2574.2011.00340.x
24. Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol.* 1996 Aug;49(8):907-16. PubMed PMID: 8699212. https://doi.org/10.1016/0895-4356(96)00025-x
25. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. *Biom J.* 2018 May;60(3):431-49. PubMed PMID: 29292533. PMCID: PMC5969114. Epub 20180102. https://doi.org/10.1002/bimj.201700067
26. Groenwold RHH, Shofty I, Miočević M, van Smeden M, Klugkist I. Adjustment for unmeasured confounding through informative priors for the confounder-outcome relation. BMC Med Res Methodol. 2018 Dec 22;18(1):174. PubMed PMID: 30577773. PMCID: PMC6303957. Epub 20181222. https://doi.org/10.1186/s12874-018-0634-3
27. Altman DG, Bland JM. How to obtain the P value from a confidence interval. *Bmj*. 2011;343:d2304. PubMed PMID: 22803193. https://doi.org/10.1136/bmj.d2304
28. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol.* 2008 Apr;61(4):344-9. PubMed PMID: 18313558. https://doi.org/10.1016/j.jclinepi.2007.11.008
29. Bunsiripaiboon P, Sornmayura P, Wilasrusmee C, Lertsithichai P. The prognostic significance of microvessel density in intrahepatic cholangiocarcinoma. *J Med Assoc Thai.* 2010 Jan;93(1):66-72. PubMed PMID: 20196413. http://www.mat.or.th/journal
30. Tamandl D, Herberger B, Gruenberger B, Puhalla H, Klinger M, Gruenberger T. Influence of hepatic resection margin on recurrence and survival in intrahepatic cholangiocarcinoma. Ann Surg Oncol. 2008 Oct;15(10):2787-94. PubMed PMID: 18685896. Epub 20080807. https://doi.org/10.1245/s10434-008-0081-1
31. Fisher SB, Patel SH, Kooby DA, Weber S, Bloomston M, Cho C, Hatzaras I, Schmidt C, Winslow E, Staley CA 3rd, Maithel SK. Lymphovascular and perineural invasion as selection criteria for adjuvant therapy in intrahepatic cholangiocarcinoma: a multi-institution analysis. HPB (Oxford). 2012 Aug;14(8):514-22. PubMed PMID: 22762399. PMCID: PMC3406348. Epub 20120522. https://doi.org/10.1111/j.1477-2574.2012.00489.x
32. Moris D, Palta M, Kim C, Allen PJ, Morse MA, Lidsky ME. Advances in the treatment of intrahepatic cholangiocarcinoma: An overview of the current and future therapeutic landscape for clinicians. *CA Cancer J Clin.* 2023 Mar;73(2):198-222. PubMed PMID: 36260350. Epub 20221019. https://doi.org/10.3322/caac.21759
33. Mitani AA, Haneuse S. Small Data Challenges of Studying Rare Diseases. JAMA Netw Open. 2020 Mar 2;3(3):e201965. PubMed PMID: 32202640. Epub 20200302. https://doi.org/10.1001/jamanetworkopen.2020.1965
34. Gagne JJ, Thompson L, O'Keefe K, Kesselheim AS. Innovative research methods for studying treatments for rare diseases: methodological review. Bmj. 2014 Nov 24;349:g6802. PubMed PMID: 25422272. PMCID: PMC4242670. Epub 20141124. https://doi.org/10.1136/bmj.g6802
35. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond "p< 0.05". Taylor & Francis; 2019. p. 1-19. https://doi.org/10.1080/00031305.2019.1583913
36. Goodman SN. How sure are you of your result? Put a number on it. Nature. 564. England2018. p. 7. https://doi.org/10.1038/d41586-018-07589-2
37. Stensrud MJ, Hernán MA. Why Test for Proportional Hazards? *Jama*. 2020 Apr 14;323(14):1401-2. PubMed PMID: 32167523. PMCID: PMC11983487. https://doi.org/10.1001/jama.2020.1267
38. Halpern SD, Karlawish JH, Berlin JA. The continuing unethical conduct of underpowered clinical trials. *Jama*. 2002 Jul 17;288(3):358-62. PubMed PMID: 12117401. https://doi.org/10.1001/jama.288.3.358

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Original Article DOI: https://doi.org/10.23950/jcmk/17474

Understanding the Healthcare Landscape for Pediatric Epilepsy Patients in Kazakhstan: Challenges and Opportunities

Altynshash Jaxybayeva[1], Dina Kalinina[2]

[1]*Department of Neurology, Astana Medical University, Astana, Kazakhstan*
[2]*Department of Medicine, Nazarbayev University School of Medicine, Astana, Kazakhstan*

Received: 2025-08-22.
Accepted: 2025-11-27.

J Clin Med Kaz 2025; 22(6): 103–108

Corresponding author:
Dina Kalinina.
E-mail: dina.kalinina@nu.edu.kz.
ORCID: 0000-0003-3887-1170.

ABSTRACT

Background: Epilepsy is a chronic neurological disorder affecting millions worldwide, with a particularly high burden in low- and middle-income countries (LMICs). In Kazakhstan, a middle-income country, the epidemiological situation regarding pediatric epilepsy remains underexplored, especially in terms of treatment availability and healthcare resources.

Objective: This study aimed to assess the regional distribution of epilepsy prevalence and the availability of healthcare resources for pediatric epilepsy patients across Kazakhstan.

Methods: Using secondary data collected from regional healthcare reports 2023, we analyzed the prevalence of epilepsy, the distribution of epilepsy forms, the availability of pediatric neurologists, and the provision of antiepileptic drugs (AEDs) across 16 regions.

Results: We found significant regional disparities in epilepsy prevalence, ranging from 54.20 to 4,263.86 per 100,000 population, with the highest prevalence in the Zhambyl and Almaty regions. The availability of pediatric neurologists varied widely, from as few as 2 specialists in some regions to 156 in Almaty. Carbamazepine and valproic acid were the most widely available AEDs, but there were gaps in the provision of other drugs, such as Diazepam and Oxcarbazepine.

Conclusion: Our findings highlight critical gaps in healthcare infrastructure and access to treatment, particularly in rural regions, which may contribute to the high treatment gap for epilepsy in Kazakhstan. The study underscores the need for targeted healthcare interventions, increased resources for pediatric epilepsy care, and policy initiatives to address regional disparities in treatment and diagnostic services.

Keywords: Epilepsy; Pediatrics; Health Resources; Anticonvulsants; Kazakhstan

Introduction

Epilepsy is a chronic neurological disease and, based on the International League Against Epilepsy (ILAE) report, should be defined by at least two unprovoked seizures occurring more than 24 hours apart, or one unprovoked seizure and a probability of subsequent seizures with a risk of at least 60% of the general recurrence risk within ten years of the first two unprovoked seizures, or the diagnosis of epilepsy syndrome [1]. According to Meyer et al. (2009) and Shan et al. (2024), approximately 50 million people worldwide suffer from epilepsy (PWE), of which 25% or 10.5 million are children under the age of 15 [2–4]. The epilepsy burden is also higher in low- and middle-income countries (LMIC), where about 80% of PWE live [5].

Furthermore, there is a significant gap in the allocation of healthcare resources represented by a high treatment gap (75–90%) fueled by severe shortages of pediatric neurologists and limited access to diagnostic

tools like EEG and MRI [6–8]. Additionally, due to limited resources in the LMIC, most epilepsy cases associated with genetic abnormalities remain underdiagnosed [9–11].

The epidemiological situation with epilepsy in Kazakhstan, which is also classified as a middle-income country, remains unclear, especially in the context of pediatric patients. A recent study analyzed epilepsy-related data for ICD-10 codes (G40.0-G40.9) from the Unified National Electronic Health System of Kazakhstan over a seven-year span and revealed growing incidence and prevalence rates from 26.15 in 2014 to 88.80 in 2020 and 26.06 in 2014 to 73.10 in 2020 per 100,000 people, respectively [12]. However, the studies still lack clarity on the availability of treatment and diagnostic resources.

In the context of the work and duties of the chief neurologist of Kazakhstan, in 2023 the reports about the population under 18 years old, the total number of epilepsy cases, the number of pediatric neurologists, and drug availability from regions were received. This study aims to analyze the regional reports-based data and shed light on the current healthcare resources and challenges faced by pediatric epilepsy patients across various regions in Kazakhstan.

Methods

Data sources

This is a descriptive, observational study that uses secondary data analysis methods. Data on the population under 18 years old, the total number of epilepsy cases, the number of pediatric neurologists, and drug availability for the 16 regions (Akmola region, Aktobe region, Almaty, Almaty region, Astana, Atyrau region, East Kazakhstan region, Karaganda region, Kostanay region, Kyzylorda region, Mangystau region, North Kazakhstan region, Pavlodar region, Turkestan region, West Kazakhstan region, and Zhambyl region) in Kazakhstan were obtained in 2023. The chief neurologist of Kazakhstan collected data from regional healthcare reports to assess the epidemiological situation of epilepsy and the availability of healthcare resources for pediatric epilepsy patients.

Statistical analysis

We conducted a descriptive epidemiological analysis to evaluate the regional distribution of epilepsy prevalence across Kazakhstan. The prevalence rates were calculated by dividing the total number of epilepsy cases by the regional pediatric population (under 18 years old) and standardizing the results per 100,000 population. The total number of epilepsy cases was analyzed by age groups (0-5, 6-18 years old) and regions. Data were disaggregated by age and geographical location to identify patterns in disease burden. Age-specific case distributions were calculated to assess variations across the pediatric population and geographic regions. The distribution of epilepsy forms across Kazakhstan was examined. We categorized the data by specific forms of epilepsy and analyzed regional variations to understand the prevalence of different types of epilepsy. Log-scale transformation was applied to the total number of epilepsy cases for each form of epilepsy. The log transformation was performed by applying the natural logarithm (log) to the total number of epilepsy cases for each region and each form of epilepsy. This approach reduced the impact of extreme values, allowing for more effective comparisons across regions with large discrepancies in case numbers. All analyses were performed using the R software (version 4.3.0).

Results

Prevalence of epilepsy cases among children in the Kazakhstan region

Epilepsy prevalence ranges from 54.20 to 4,263.86 per 100,000 population. The regions with the highest prevalence rates are Zhambyl (4,263.86 per 100,000) and the Almaty region (1,417.14 per 100,000). Other regions like East Kazakhstan (399.25 per 100,000), Karaganda (295.33 per 100,000), and Akmola (282.07 per 100,000) show moderate epilepsy prevalence rates. Regions like Mangystau (54.20 per 100,000), West Kazakhstan (185.54 per 100,000), and Pavlodar (94.23 per 100,000) have lower rates (Figure 1).

Total cases of epilepsy by age and region

The total number of epilepsy cases varied from 142 to 2,363. The highest values were observed in the Turkestan, Almaty, and Zhambyl regions, with 2,363, 1,815, and 1,636 cases, respectively. Other regions with a considerable number of cases included Astana (792 cases), the Akmola region (679 cases), and East Kazakhstan (723 cases). In contrast, regions such as the Atyrau region (142 cases), West Kazakhstan (524 cases), and the Pavlodar region (579 cases) reported fewer cases. In regard to age distribution, in all regions, epilepsy cases were most common among 6-18-year-olds (Figure 2).

Distribution of Epilepsy Forms in Kazakhstan

The log-scaled number of cases for each epilepsy form showed that epilepsy (15,126 cases) was the most distributed form in Kazakhstan. While other rare epilepsy syndromes are less represented, children are more often diagnosed with Dravet (245 cases) and West syndrome (187 cases), and only 8 cases were registered with Ohtahara syndrome (Figure 3).

Distribution of pediatric neurologists across different regions of Kazakhstan

The number of pediatric neurologists across different regions of Kazakhstan varied from 2 to 156 specialists (Figure 4). Almaty has the highest number of pediatric neurologists (156); West Kazakhstan (44), Astana (39) and Shymkent (39) also have significant numbers. Regions like Pavlodar (4), Mangystau (6), and Kyzylorda (8) have very low numbers of pediatric neurologists, and the Almaty region is presented with only 2 specialists.

Distribution of antiepileptic drugs across regions of Kazakhstan

Valproic acid is the most commonly available antiepileptic drug (AED) in most regions with the highest numbers reported in the Turkestan region (1,738 patients), Almaty region (1,118 patients), and Shymkent (803 patients) (Figure 5). Carbamazepine also showed high provision rates, especially in the Almaty region (378 patients), Shymkent (340 patients), and Turkestan region (300 patients). Similarly. Diazepam and oxcarbazepine had relatively lower patient numbers, with some regions showing no patients receiving these drugs. For example, West Kazakhstan had 26 patients not provided with Diazepam. Almaty, Aktobe, East Kazakhstan, Karaganda, Kostanay, North Kazakhstan, Mangystau, and Turkestan regions reported zero patients not provided with any AED.

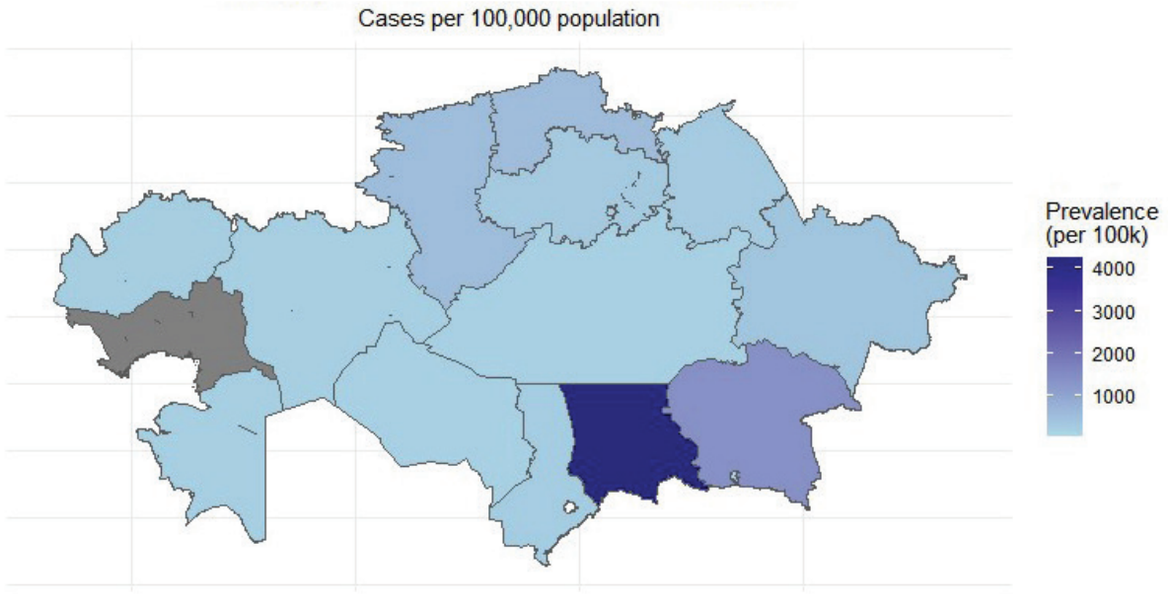


Figure 1 – Map illustrating Epilepsy Prevalence among Children in Regions of Kazakhstan

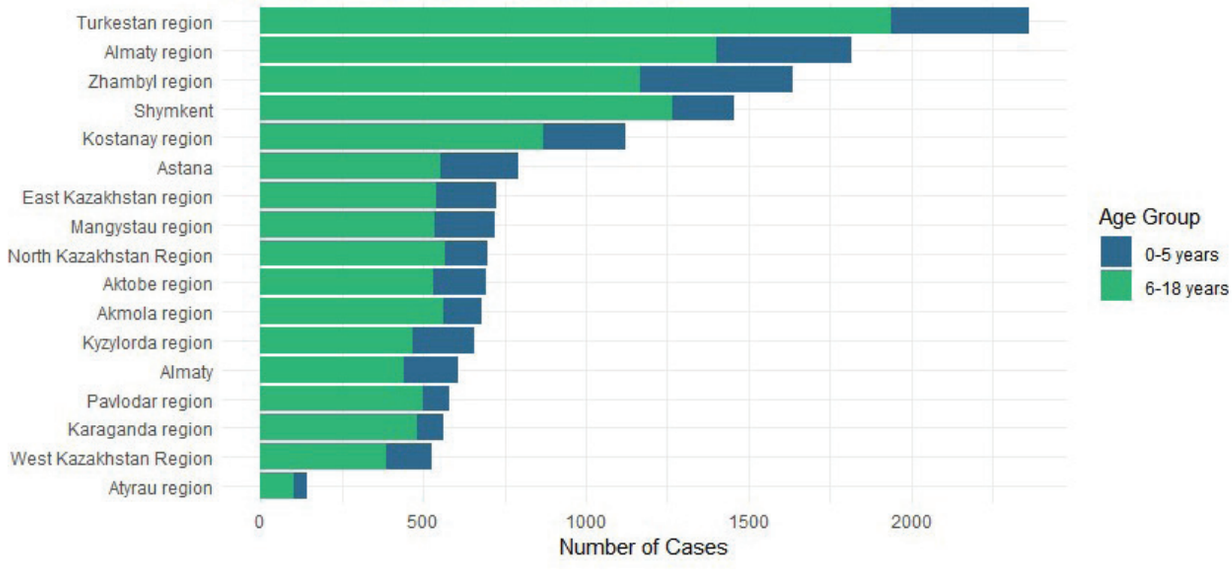


Figure 2 – Total Cases of All Epilepsy Forms Among Children by Regions and Age Groups

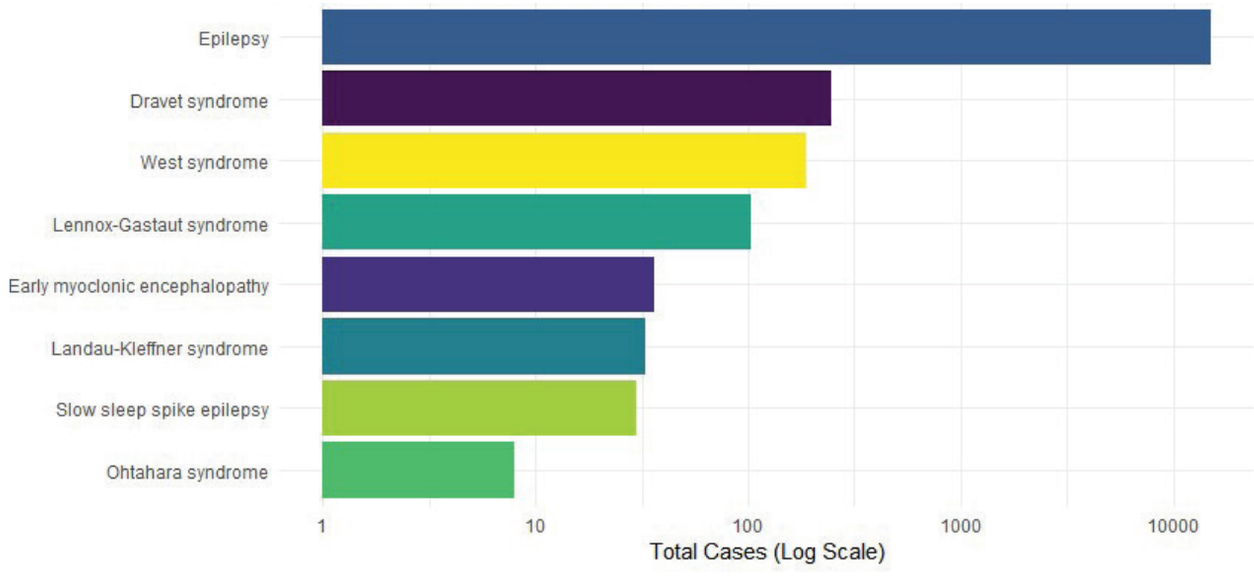


Figure 3 – Log-scaled Distribution of Epilepsy Forms in Kazakhstan

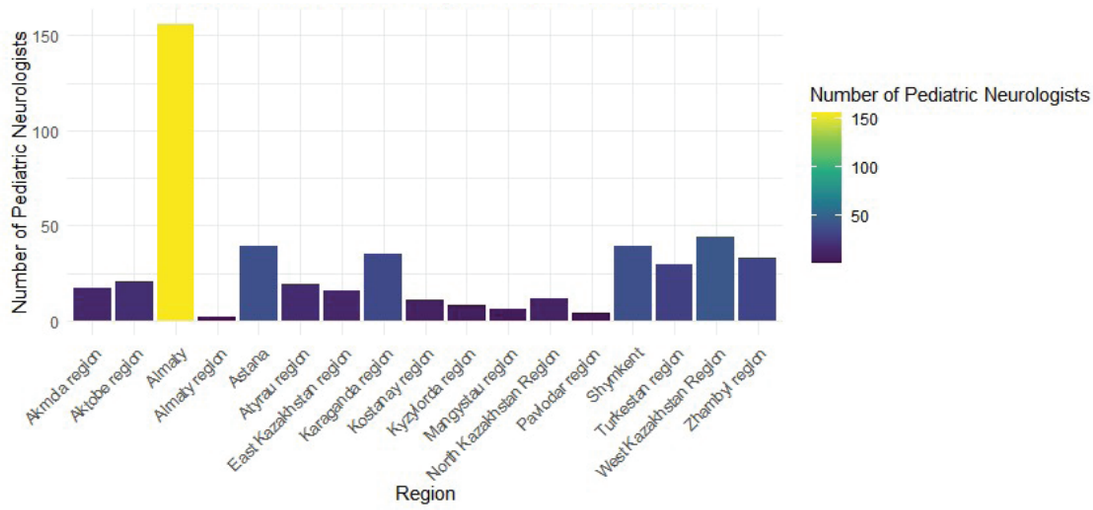


Figure 4 – Number of Pediatric Neurologists in Different Regions

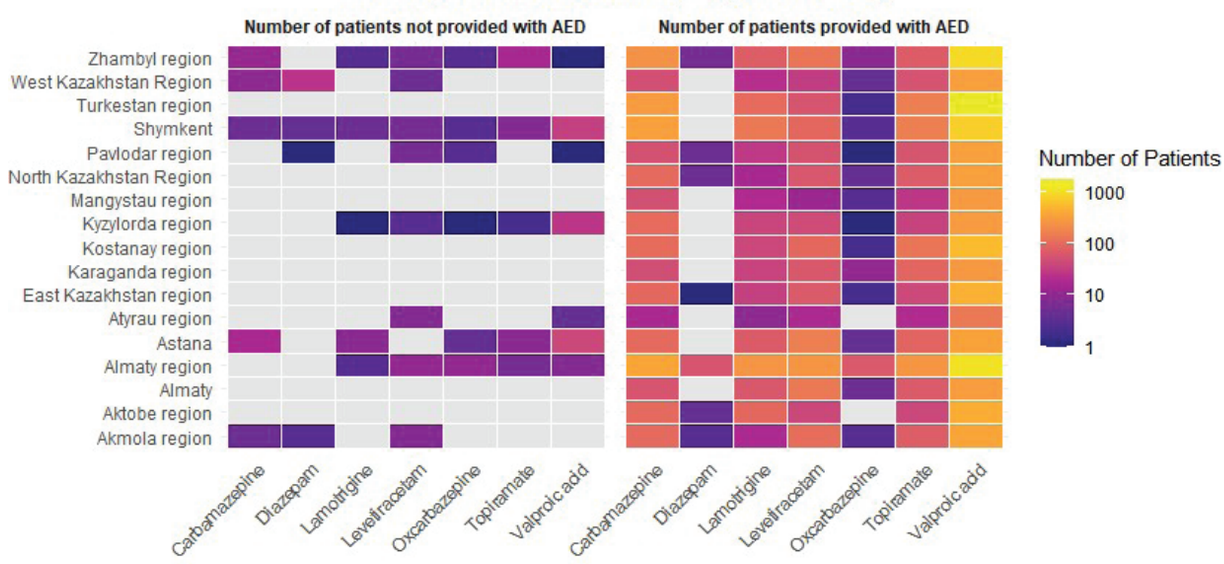


Figure 5 – Heat Map of Antiepileptic Drug Provision by Region and Drugs

Grey color – 0 patients, gradient colors varied based on the log-scaled number of patients. AED – antiepileptic drugs.

Discussion

Our study revealed a wide variation in the epilepsy prevalence rates in children and the availability of healthcare resources across different regions of Kazakhstan. The regions with the highest prevalence rates (Zhambyl and Almaty region) and the lowest (Mangystau and West Kazakhstan) could potentially show correlations with healthcare access, regional healthcare infrastructure, socioeconomic status, and population demographics. There are also significant regional differences in the number of pediatric neurologists, with some regions having a very low number (Mangystau and Pavlodar), while others like Almaty and Astana have a relatively higher concentration of specialists.

Since independence, the national health system of Kazakhstan has undergone many reforms, but access to quality healthcare is still uneven [13]. For example, in some cities such as Astana and Almaty, the health infrastructure, including the medical workforce and diagnostic equipment, is well developed, while in rural areas, these resources are limited, leading to delayed diagnoses and insufficient care [14]. The observed highly varied distribution of pediatric neurologists across various regions aligns with a study that revealed differences in the median number of neurologists per 100,000 people between low-income countries (0.03) and high-income countries (2.96) [15].

The availability of AEDs in Kazakhstan may also reflect trends in LMICs globally. Despite their relative ineffectiveness and higher adverse effects, old-generation cheap drugs like carbamazepine and valproic acid remain more prevalent due to their cost and availability [16,17]. Similar patterns have been described in tertiary-care settings, where older agents continue to dominate treatment because of affordability and accessibility. For instance, a study comparing conventional and newer ASMs in India found that conventional drugs were still widely prescribed despite their less favorable safety profile, mainly due to their lower cost and availability [18]. This mirrors the persistent reliance on older ASMs seen in Kazakhstan. Unfortunately, some AEDs (diazepam) remain unavailable for children with

epilepsy in some regions of Kazakhstan. Other countries also experience limited resources in drug provision, where 67% of all PWE may not have any AEDs [19].

Treatment disparities may be further compounded by differences in diagnostic capacity. In Kazakhstan, access to EEG and neuroimaging is uneven, and this may influence both treatment choices and long-term outcomes. Evidence from clinical studies suggests that delays in appropriate treatment or incomplete diagnostic workup can worsen disease trajectories. For example, Dossov (2025) showed that in pharmacoresistant focal epilepsy, longer disease duration was associated with more intensive intraoperative management needs, indirectly emphasizing the importance of early diagnosis and optimized therapy—both of which depend on adequate regional resources [20]. Genetic testing and advanced molecular diagnostics remain limited as well, despite growing evidence of their importance. A recent genomic investigation of infantile and childhood epileptic encephalopathies in Kazakhstan highlighted significant gaps in access to genetic evaluation and emphasized its urgent priority for improving early diagnosis and individualized care [21]. Together, these findings show that incomplete diagnostic pathways can delay appropriate treatment and contribute to worse clinical outcomes.

To address the disparities in pediatric epilepsy care across Kazakhstan, strategic policy initiatives should focus on expanding the healthcare workforce, particularly by increasing the number of pediatric neurologists in underserved regions. Targeted recruitment, incentives for specialists to practice in resource-limited areas, and enhanced access to AEDs can achieve this [22]. Additionally, investments in diagnostic infrastructure, such as EEG, MRI, and CT equipment, are essential to improve the early and accurate diagnosis of epilepsy [23]. Strengthening diagnostic pathways is especially relevant in light of recent work demonstrating the value of structured diagnostic approaches and predictive tools—for example, machine-learning-based risk stratification models for severe pediatric epileptic conditions such as status epilepticus, which showed promising accuracy in outcome prediction [24]. Although such tools are not yet widely used in Kazakhstan, improving access to diagnostics would make their implementation feasible in the future.

Training and education programs for healthcare providers are equally crucial, especially in regions with limited expertise [25]. These initiatives could include continuing medical education programs, telemedicine support to connect rural practitioners with specialists in urban centers, and collaborations with international organizations to facilitate knowledge exchange and adoption of best practices [26,27]. By integrating these strategies into a cohesive national policy, Kazakhstan can take significant steps toward ensuring equitable and effective care for all children with epilepsy.

Limitations

The study relies on secondary data from regional healthcare reports. Discrepancies or potential biases in data reporting could affect the accuracy of the findings. Some regions may underreport epilepsy cases due to limited diagnostic resources or a lack of standardized reporting systems. Additionally, as a cross-sectional study, this analysis provides a snapshot of the current situation but does not capture temporal trends or changes in healthcare resource allocation over time. Furthermore, our analysis did not fully explore the impact of urban-rural disparities or socio-demographic barriers on the access to care for pediatric epilepsy patients.

Conclusion

This study highlights the significant disparities in the prevalence of pediatric epilepsy and the distribution of healthcare resources across the regions of Kazakhstan. These findings emphasize the urgent need to address healthcare inequities, particularly in resource-limited regions. Increasing the availability of trained pediatric neurologists, diagnostic tools, and AEDs is essential to ensure equitable care for children with epilepsy across Kazakhstan. Future studies should explore the underlying causes of regional disparities and assess the impact of socioeconomic, cultural, and systemic factors on epilepsy management. This evidence can guide targeted interventions and inform policies aimed at reducing the epilepsy treatment gap and improving outcomes for pediatric patients nationwide.

Author Contributions: Conceptualization, A.J. and D.K.; methodology, D.K.; validation, A.J.; formal analysis, D.K.; investigation, A.J.; resources, A.J.; data curation, A.J.; writing – original draft preparation, D.K.; writing – review and editing, D.K. and A.J.; visualization, D.K.; supervision, A.J.; project administration, not applicable; funding acquisition, not applicable. Both authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

Ethics approval consent to participate: This study was approved by the Institutional Review Ethics Committee of the NROC (NROC-IREC – 034, July 30, 2025). Given the retrospective nature of the study and the fact that complete anonymization of all patient data was ensured, the Institutional Review Ethics Committee waived the requirement for informed consent.

References

1. Fisher, R.S.; Acevedo, C.; Arzimanoglou, A.; Bogacz, A.; Cross, J.H.; Elger, C.E.; Engel, J.; Forsgren, L.; French, J.A.; Glynn, M.; et al. ILAE Official Report: A Practical Clinical Definition of Epilepsy. Epilepsia 2014, 55, 475–482, https://doi.org/10.1111/epi.12550.
2. Meyer, A.-C.; Dua, T.; Ma, J.; Saxena, S.; Birbeck, G. Global Disparities in the Epilepsy Treatment Gap: A Systematic Review. Bull. World Health Organ. 2010, 88, 260–266, https://doi.org/10.2471/BLT.09.064147.
3. Shan, T.; Zhu, Y.; Fan, H.; Liu, Z.; Xie, J.; Li, M.; Jing, S. Global, Regional, and National Time Trends in the Burden of Epilepsy, 1990–2019: An Age-Period-Cohort Analysis for the Global Burden of Disease 2019 Study. Front. *Neurol*. 2024, 15, 1418926, https://doi.org/»0.3389/fneur.2024.1418926.
4. Guerrini, R. Epilepsy in Children. *The Lancet* 2006, 367, 499–524, https://doi.org/10.1016/S0140-6736(06)68182-8.
5. Leonardi, M.; Ustun, T.B. The Global Burden of Epilepsy. Epilepsia 2002, 43, 21–25, https://doi.org/10.1046/j.1528-1157.43.s.6.11.x.
6. Mbuba, C.K.; Ngugi, A.K.; Newton, C.R.; Carter, J.A. The Epilepsy Treatment Gap in Developing Countries: A Systematic Review of the Magnitude, Causes, and Intervention Strategies. Epilepsia 2008, 49, 1491–1503, https://doi.org/10.1111/j.1528-1167.2008.01693.x.
7. Wilmshurst, J.M.; Birbeck, G.L.; Newton, C.R. Epilepsy Is Ubiquitous, but More Devastating in the Poorer Regions of the World… or Is It? Epilepsia 2014, 55, 1322–1325, https://doi.org/10.1111/epi.12602.
8. Di Noia, S.; Bonezzi, L.; Accorinti, I.; Bartolini, E. Diagnosis and Classification of Pediatric Epilepsy in Sub-Saharan Africa: A Comprehensive Review. *JCM* 2024, 13, 6396, https://doi.org/10.3390/jcm13216396.
9. Kaiyrzhanov, R.; Zharkinbekova, N.; Guliyeva, U.; Ganieva, M.; Tavadyan, Z.; Gachechiladze, T.; Salayev, K.; Guliyeva, S.; Isayan, M.; Kekenadze, M.; et al. Elucidating the Genomic Basis of Rare Pediatric Neurological Diseases in Central Asia and Transcaucasia. Nat Genet 2024, 56, 2582–2584, https://doi.org/10.1038/s41588-024-02016-x.
10. Essajee, F.; Urban, M.; Smit, L.; Wilmshurst, J.M.; Solomons, R.; Van Toorn, R.; Moosa, S. Utility of Genetic Testing in Children with Developmental and Epileptic Encephalopathy (DEE) at a Tertiary Hospital in South Africa: A Prospective Study. Seizure: *European Journal of Epilepsy* 2022, 101, 197–204, https://doi.org/10.1016/j.seizure.2022.09.001.
11. Esterhuizen, A.I.; Tiffin, N.; Riordan, G.; Wessels, M.; Burman, R.J.; Aziz, M.C.; Calhoun, J.D.; Gunti, J.; Amiri, E.E.; Ramamurthy, A.; et al. Precision Medicine for Developmental and Epileptic Encephalopathies in Africa—Strategies for a Resource-Limited Setting. Genetics in Medicine 2023, 25, 100333, https://doi.org/10.1016/j.gim.2022.11.002.
12. Akhmedullin, R.; Kozhobekova, B.; Gusmanov, A.; Aimyshev, T.; Utebekov, Z.; Kyrgyzbay, G.; Shpekov, A.; Gaipov, A. Epilepsy Trends in Kazakhstan: A Retrospective Longitudinal Study Using Data from Unified National Electronic Health System 2014–2020. Seizure: European Journal of Epilepsy 2024, 122, 58–63, https://doi.org/10.1016/j.seizure.2024.09.022.
13. Makhanov, K. Urbanization and Spatial Demographic Trends in Kazakhstan 2021.
14. Shaltynov, A.; Raushanova, A.; Jamedinova, U.; Sepbossynova, A.; Myssayev, A.; Myssayev, A. Health-Care Accessibility Assessment in Kazakhstan. Open Access Maced J Med Sci 2021, 9, 89–94, https://doi.org/10.3889/oamjms.2021.5704.
15. Trinka, E.; Kwan, P.; Lee, B.; Dash, A. Epilepsy in Asia: Disease Burden, Management Barriers, and Challenges. Epilepsia 2019, 60, 7–21, https://doi.org/10.1111/epi.14458.
16. Bankole, N.D.A.; Dokponou, Y.C.H.; De Koning, R.; Dalle, D.U.; Kesici, Ö.; Egu, C.; Ikwuegbuenyi, C.; Adegboyega, G.; Yang Ooi, S.Z.; Dada, O.E.; et al. Epilepsy Care and Outcome in Low- and Middle-Income Countries: A Scoping Review. JNRP 2024, 15, 8–15, https://doi.org/10.25259/JNRP_527_2023.
17. Meinardi, H.; Scott, R.A.; Reis, R.; On Behalf Of The Ilae Commission On The Developing World, J.W.A.S.S. The Treatment Gap in Epilepsy: The Current Situation and Ways Forward. Epilepsia 2001, 42, 136–149, https://doi.org/10.1046/j.1528-1157.2001.32800.x.
18. Thomas, M.; Badyal, D.K.; Pandian, J.D. Comparative Effectiveness-Safety of Conventional versus Newer Antiepileptics in Epileptic Patients in a Tertiary Care Hospital, India. J CLIN MED KAZ 2023, 20, 50–54, https://doi.org/10.23950/jcmk/13909.
19. Asadi-Pooya, A.A. Availability of Antiepileptic Drugs: Politicians' Roles. Epilepsia 2016, 57, 671–671, https://doi.org/10.1111/epi.13323.
20. Dossov, M. The Influence of Disease Duration in Pharmacoresistant Focal Epilepsy and Type of Antiepileptic Therapy on Intraoperative Infusion Therapy Parameters. J CLIN MED KAZ 2025.
21. Jaxybayeva, A.; Nauryzbayeva, A.; Khamzina, A.; Takhanova, M.; Abilhadirova, A.; Rybalko, A.; Jamanbekova, K. Genomic Investigation of Infantile and Childhood Epileptic Encephalopathies in Kazakhstan: An Urgent Priority. Front. Neurol. 2021, 12, https://doi.org/10.3389/fneur.2021.639317.
22. Pellinen, J. Treatment Gaps in Epilepsy. Front. *Epidemiol*. 2022, 2, 976039, https://doi.org/10.3389/fepid.2022.976039.
23. Government of Kazakhstan Official Information Source of the Prime Minister of the Republic of Kazakhstan | Children's Rehabilitation: What Will Change after Implementing Compulsory Social Health Insurance Available online: https://primeminister.kz/en/news/reviews/detskaya-reabilitaciya-chto-izmenitsya-posle-vnedreniya-osms (accessed on 20 August 2025).
24. Rushitha, A.L.; K, D.; K, R.; Periasamy, P.; Choudhary, A.K. Diagnostic Evaluation of the STEPSS Score and Machine Learning Models in Predicting Outcomes in Pediatric Status Epilepticus: A Prospective Observational Study. *J CLIN MED KAZ* 2025.
25. Briscoeabath, C. The Harsh Reality of Disparities in Pediatric Epilepsy Surgery. Pediatric Epilepsy Surgery Alliance 2024.
26. Owens, S.; Sirven, J.I.; Shafer, P.O.; Fishman, J.; Wild, I.; Findley, M.; Derry, R.; Walters, J.; Kopplin, V.; Kakacek, J. Innovative Approaches Reaching Underserved and Rural Communities to Improve Epilepsy Care: A Review of the Methodology of the Connectors Project. Epilepsy & Behavior 2019, 90, 273–283, https://doi.org/10.1016/j.yebeh.2018.09.029.
27. WHO Telemedicine in Kazakhstan: Smart Health Services Delivery Available online: https://www.who.int/europe/news-room/photo-stories/item/telemedicine-in-kazakhstan-smart-health-services-delivery (accessed on 20 August 2025).

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Review Article DOI: https://doi.org/10.23950/jcmk/17362

Mucosal Immune Markers and Their Impact on Antimicrobial Resistance: A Systematic Review

Olabisi Promise Lawal[1], Aliyu Olanrewaju Olaniyi[2], Uche Cyprain Tochukwu[3], Catherine Rono[4], Mayowa Josiah Badmus[5], Misbahudeen Olohuntoyin Abdulkareem[6], Udeichi Mmesoma Odinaka[7], Idris Olumide Orenolu[8], Udeichi Nnenna Chidimma[9]

[1]*Department of Medical Laboratory Science, University of Benin, Benin City, Nigeria*
[2]*Stockport NHS foundation trust, Stepping Hill Hospital, Stockport, Manchester. United Kingdom.*
[3]*Department of Epidemiology and Evidence-based Medicine. First Moscow State Medical University named after I. M. Sechenov, Moscow, Russia.*
[4]*Department of Immunology; Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA*
[5]*Faculty of Dentistry, Lagos University Teaching Hospital, Lagos, Nigeria.*
[6]*Department of Internal Medicine, Kwara State University Teaching Hospital, Ilorin, Kwara State. Nigeria*
[7]*Centre For Infectious Diseases, 68 Nigerian Army Reference Hospital, Yaba.*
Lagos, Nigeria
[8]*Infection Prevention and Control, Saskatchewan Health Authority, Saskatchewan, Canada*
[9]*Critical care department, Iwosan Lagoon Hospital, Ikoyi*

Received: 2025-07-01.
Accepted: 2025-09-07.

J Clin Med Kaz 2025; 22(6): 109–116

Corresponding author:
Olabisi Promise Lawal .
E-mail: lawalolabisipromise@gmail.com.
ORCID: 0009-0008-5287-5999.

Abstract

Objectives: This systematic review aimed to evaluate the role of secretory immunoglobulin A (sIgA) and related mucosal immune markers in modulating susceptibility to colonization or infection by antimicrobial-resistant (AMR) bacteria in human and animal models.

Methods: Studies were included if they assessed mucosal sIgA levels about AMR outcomes, reported original data, and were published in English between 2015 and June 2025. Exclusion criteria included a lack of mucosal immune data or AMR-related endpoints. Databases searched included PubMed, ScienceDirect, and EBSCO (Knowledge and Library Hub), with the last search performed on June 10, 2025. The quality of the included studies was assessed using the checklist for quantitative studies, and data were extracted and synthesized narratively, stratified by immune markers, study design, and mucosal compartment.

Results: After the systematic search of the databases considered for this review and review of available literature indexed in them, eleven studies that met the inclusion criteria of this study were included in this review. The mucosal sites investigated in most of these studies were gastrointestinal, respiratory, oral, and genitourinary mucosa. Of the eleven (11) studies, four carried out investigations in humans while the remaining seven utilized animal models. All the included studies reported secretory immunoglobulin A (sIgA) level, whereas five of the studies reported other indicators, such as cytokines and IgG, alongside sIgA levels. There was consistency in their results as increased sIgA levels were linked to decreased infection or colonization by AMR pathogens. Moreover, vaccination-based interventional studies showed an increase in sIgA post-vaccination or probiotic treatment.

Conclusion: Although in this review we included studies with varying methodological approaches as well as studies featuring humans and animal models, the evidence from this review indicates that mucosal sIgA offers protection against antimicrobial-resistant (AMR) bacteria colonization. Therefore, findings from this study suggest that in the quest for addressing AMR that approaches targeting mucosal immunity preventative measures should be explored through standardized trials.

Keywords: Mucosal Immune Markers; Secretory IgA; Mucosal Immunity; Antimicrobial Resistance; Host-Pathogen Interaction; Colonization Resistance.

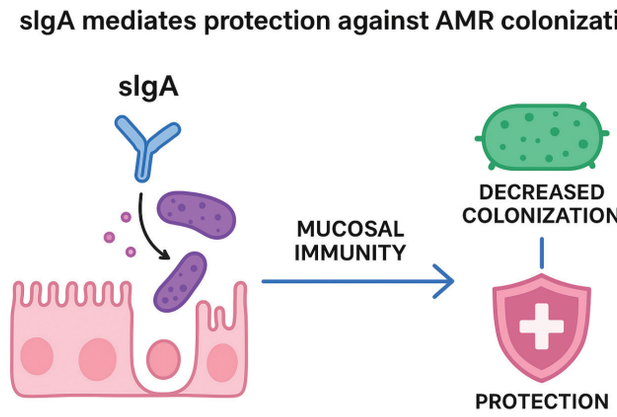


Figure 1 – Graphical Abstract: sIgA-mediated protection against AMR colonization.

Introduction

Antimicrobial resistance (AMR) has emerged as a major public health concern, jeopardizing the efficacy of life-saving antimicrobials and challenging modern medicine practices. According to the World Health Organization (WHO), AMR is now one of the top ten global public health threats to humanity, with serious implications for common infection treatment, surgical procedure safety, and the viability of immunosuppressive therapies like chemotherapy and organ transplantation [1]. In 2019, an estimated 4.95 million deaths were linked to bacterial AMR, including 1.27 million deaths directly attributed to resistant infections [1, 2].

Without prompt and coordinated global actions, the yearly death toll from AMR might reach 10 million by 2050 [3, 4]. The situation has been compounded by decades of antibiotic overuse in both the human and animal fields, as well as inadequate investment in the discovery of novel antimicrobials [4]. As recently highlighted in United Nations debates, AMR poses an existential danger not just to global health but also to food security, economic stability, and national security [5].

This rising issue has heightened interest in alternative methods of defense against infectious microorganisms, particularly the host's immune system. Mucosal immunity, which is dominated by secretory immunoglobulin A (sIgA) and supported by cytokines, innate cells, and the microbiota, is at the forefront of avoiding microbial colonization at mucosal surfaces, where most infections occur [6]. Mucosal areas such as the gastrointestinal, pulmonary, and genitourinary tracts are crucial interfaces between host and environment, making them ideal battlegrounds in the fight against AMR infections. Despite its critical function in pathogen exclusion, immune regulation, and shaping the mucosal microbiome, mucosal immunity has received relatively little attention in mainstream AMR studies. A better understanding of how mucosal immune components, notably sIgA, interact with drug-resistant pathogens could lead to novel preventive and treatments that avoid traditional antibiotic use.

The mucosal immune system is the first line of defense against antimicrobial-resistant (AMR) pathogens. Secretory immunoglobulin A (sIgA), the most common antibody isotype in mucosal secretions, provides broad protection by neutralizing pathogens, regulating the microbiota, and encouraging immunological exclusion. The polymeric nature of sIgA allows for multivalent binding to antigens, which promotes pathogen clearance via mucus trapping [7,8]. These functions are executed through specific mechanisms, such as transcytosis via the polymeric immunoglobulin receptor (pIgR), which allows sIgA to be secreted over epithelial barriers into mucosal regions [8]. Human studies, for example, have shown that increased intestinal sIgA corresponds with lower colonization by Helicobacter pylori and multidrug-resistant Enterobacteriaceae, supporting the notion that increasing mucosal antibody levels can prevent AMR pathogen settlement at the early stages [9].

The use of probiotics further explains the relationship between microbial regulation, mucosal immunity, and AMR control. In a randomized controlled trial (RCT) among children less than two years old, Lactobacillus plantarum supplementation increased salivary and fecal sIgA levels while also significantly reducing the carriage of ESBL-producing E. coli among antibiotic-treated individuals [10]. Similarly, older adults who received Bifidobacterium longum had enhanced mucosal barrier indicators and a lower prevalence of multidrug-resistant Gram-negative bacterial colonization [11-13]. Animal studies support these findings; mice colonized with probiotic strains had lower gut colonization by carbapenem-resistant Klebsiella pneumoniae, which was associated with increased sIgA production and upregulation of tight junction proteins [14, 15].

Despite the encouraging evidence, research remains disjointed and limited. Human trials examining sIgA as an endpoint in AMR colonization are uncommon and sometimes hampered by short follow-up and small sample sizes [16]. While animal models give mechanistic insight, species variations in mucosal immune architecture typically limit the direct application of findings to humans [17, 18].

These gaps can be addressed with a systematic review integrating human and animal research to determine whether variations in sIgA levels or activity at mucosal surfaces influence susceptibility to colonization or infection by AMR pathogens. This review evaluates evidence on the role of secretory immunoglobulin A (sIgA) in antimicrobial resistance (AMR), focusing on how variations in sIgA levels at mucosal sites influence host susceptibility to resistant pathogens. It also identifies methodological gaps and highlights strategies—such as mucosal vaccines and probiotic-based interventions—that could leverage mucosal immunity in the fight against AMR.

Methods

Study Question

In carrying out this review, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines in carrying out this study [19], and the research question was structured using the PICO framework (Population, Intervention/Exposure, Comparator, Outcome) [20]. The PICO framework for this study is explained below:

Population (P): Human and animal populations exposed to or at risk of infection or colonization by antimicrobial-resistant (AMR) bacteria.

Intervention/Exposure (I): Presence or elevation of mucosal immune markers at mucosal sites, primarily secretory immunoglobulin A (sIgA).

Comparator (C): Individuals or groups with lower or baseline sIgA levels

Outcome (O): Colonization, infection, or disease caused by AMR organisms, including E. coli, Shigella, MRSA, and others.

The research question: Is colonization or infection by antimicrobial-resistant bacteria in human or animal hosts influenced by the level or activity of secretory IgA at mucosal sites?

Eligibility Criteria

Studies were eligible for inclusion if they involved human participants or animal models and reported on the relationship between secretory immunoglobulin A (sIgA) and colonization or infection with antimicrobial-resistant bacteria. Eligible study designs included both observational studies (such as cross-sectional, case-control, and cohort designs) and interventional studies, defined as those in which investigators introduced a preventive or therapeutic exposure (e.g., vaccination, probiotic administration, or mucosal immunotherapy) and subsequently measured sIgA levels about antimicrobial resistance outcomes. Interventional studies encompassed both randomized controlled trials (RCTs) and non-randomized experimental designs, provided they reported quantitative assessment of sIgA and documented outcomes such as colonization, infection, or microbial resistance profiles. To ensure biological relevance, studies were required to measure sIgA at mucosal surfaces—such as the gastrointestinal, respiratory, or genitourinary tract—using validated techniques (e.g., ELISA). Only full-text articles published in English between January 2015 and June 2025 were included.

Studies were excluded if they: (1) focused solely on systemic immunity (e.g., serum IgG or IgM), (2) lacked antimicrobial resistance outcomes, or (3) were review articles, editorials, conference abstracts, or case reports.

Table 1 Database Search Strategy

Database	Search terms	Search yield
PubMed	*(("secretory IgA"[Title/Abstract] OR "sIgA"[Title/Abstract]) AND ("antimicrobial resistance"[Title/Abstract] OR "antibiotic resistance"[Title/Abstract] OR "drug-resistant bacteria"[Title/Abstract] OR "resistance genes"[Title/Abstract])) AND ("colonization"[Title/Abstract] OR "infection"[Title/Abstract] OR "microbial infection"[Title/Abstract])*	*26*
ScienceDirect	*("secretory IgA" OR "sIgA") AND ("antimicrobial resistance" OR "antibiotic resistance" OR "drug-resistant bacteria" OR "resistance genes") AND ("colonization" OR "infection" OR "microbial infection")*	*1393*
EBSCOhost (Knowledge and Library Hub)	*("secretory IgA" OR "sIgA") AND ("antimicrobial resistance" OR "antibiotic resistance" OR "drug-resistant bacteria" OR "resistance genes") AND ("colonization" OR "infection" OR "microbial infection")*	*43*

Information Sources and Search Strategy

A structured search strategy was developed and executed across 3 major databases: PubMed, ScienceDirect, and EBSCOhost (Knowledge and Library Hub).The search strategy combined terms for secretory IgA, antimicrobial resistance, and colonization/infection outcomes, using Boolean logic and adjusted database-specific syntax to maximize both sensitivity and precision.

The search terms for the three databases were: "secretory IgA" OR "sIgA" AND "antimicrobial resistance" OR "antibiotic resistance" OR "drug-resistant bacteria" OR "resistance genes" AND "colonization" OR "infection" OR "microbial infection". Details of the search strategy are provided in Table 1 below.

Study Selection

All records retrieved after applying appropriate database filters (such as original articles and studies published in the English language) were imported into Zotero for duplicate removal. The title and abstract screening was carried out independently by two reviewers, using pre-specified eligibility criteria. Full texts of potentially relevant studies were then reviewed independently for inclusion. Disagreements at any stage were resolved through consensus.

Data Extraction and Management

The data extraction for this study was carried out using Microsoft Excel, and we captured information about study characteristics, including authors and year of publication, countries where the studies were conducted, population features, mucosal site sampled, sIgA measurement method, AMR organism involved, and key findings from the studies. This extraction was conducted by two independent reviewers, and entries were cross-checked for consistency and accuracy, and any disagreement was resolved through consensus.

Quality Assessment

The quality of the included studies was assessed using the Checklist for Assessing the Quality of Quantitative Studies [21], as adopted in Silvani et al. (2022) study [22]. This evaluation tool was chosen for its completeness and the inclusion of all criteria deemed pertinent in assessing quantitative studies. This tool uses 14 questions for assessment, and every item receives a score accordingly. For a “Yes”, a score of 2 was assigned, for a “Partial yes”, a score of 1 was assigned, and for a “No”, 0 was assigned. Therefore, the lowest possible score to assign a question was 0, while the highest score was 2. Items indicated with N/A were not considered as the questions were not considered for the study design of the study being assessed. Thus, the highest possible score from the 14 questions, if all the questions are considered for a study, is 28. The final score was calculated by adding the total scores across the relevant components and expressing them as a percentage of the theoretical maximum. The studies were classified as strong quality (>75%), moderate quality (55%-75%), and weak quality (<55%). Two reviewers completed this quality evaluation. Study quality scores that differed between reviewers were reviewed until a consensus was reached.

Ethical Considerations

This review does not involve any human participants; rather, it involves the synthesis of publicly available data; therefore, ethical approval was not required.

Results

Study Selection

A total of 1,462 records were retrieved from three databases—PubMed (26), ScienceDirect (1,393), and EBSCO (43). Following the application of database-specific filters—including date range (2015–2025), language (English), full text articles, and study type (original quantitative research), 1,383 records were excluded. Of the remaining 79 articles, 15 duplicates were removed, and 64 articles were subjected to title and abstract screening. Twenty-six articles were selected for full-text review.

Eleven studies met the inclusion criteria and were included in this review [23-33]. The remaining 15 studies were excluded for the following reasons: lack of mucosal immune marker assessment (n = 11), review article (n = 2), and no AMR-related outcomes (n = 2).

These studies include both human and animal models and focus primarily on secretory IgA (sIgA) and its role in colonization or infection with antimicrobial-resistant organisms. The full selection process is depicted in Figure 2 (PRISMA flow chart).

Characteristics of Included Studies

This review synthesized evidence from 11 studies as contained in Supplement file 1. These studies were published between 2017 and 2024, involving both human participants (n = 4) and animal models (n = 7), in order to examine the relationship between mucosal immune markers—particularly secretory immunoglobulin A (sIgA)—and colonization or infection with antimicrobial-resistant (AMR) pathogens. These studies were conducted in countries like China (n=5), India (n=2), the USA (n=2), Italy (n=1), and Ukraine (n=1).

However, among these included studies, one (1) is an observational study [26], while ten (10) employed interventional study designs, and the interventions featured include mucosal immunotherapy, vaccination and probiotics [23-25, 27-33]. The population studied in the human studies was mainly adults and patients with periodontitis or recurrent urinary tract infections. Whereas for studies involving animals, they predominantly used murine or poultry models.

When considering the mucosal site samples, the gastrointestinal tract was the most common(n = 7), followed by the genitourinary tract (n=2), then the respiratory tract (n=1), and the oral cavity (n=1). Secretory IgA was quantitatively assessed in all the studies through ELISA or immunoassays. Aside from the sIgA, some studies additionally reported other mucosal immune markers, including mucosal IgG, IL-6, and TNF-α [25, 33].

Nevertheless, these studies targeted a variety of AMR pathogens, including multidrug-resistant Salmonella, Acinetobacter baumannii [27], MRSA [23], Campylobacter jejuni [24], and others. However, bacterial colonization, infection severity, vaccine-induced protection, and mucosal immune response agnitude were the outcomes measured.

Quality Assessmentm

The total adjusted scores for the eleven studies included in this review ranged from 14 to 27, with maximum possible scores between 22 and 28, depending on the applicability of the criteria. All of the studies are of high quality except for the Gudaryan et

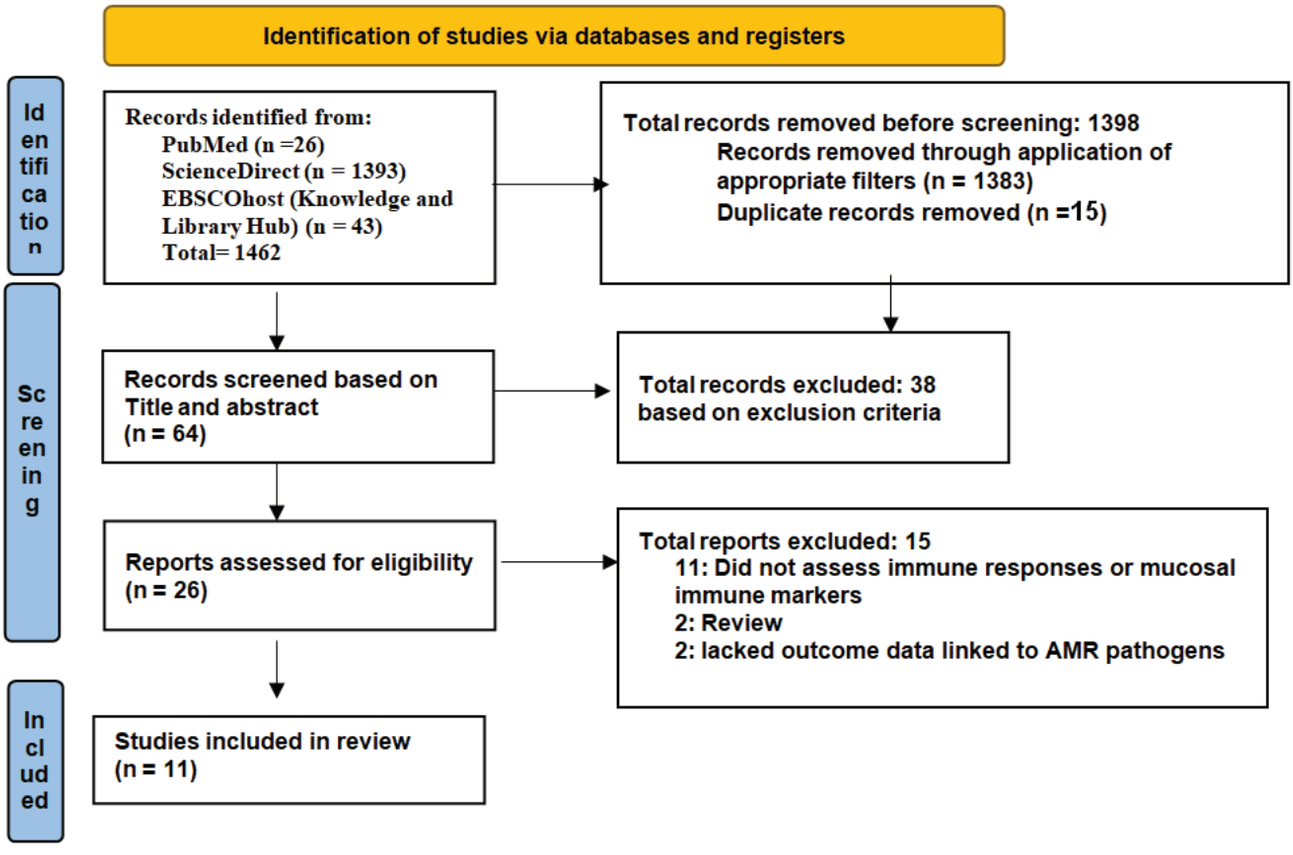


Figure 12– Prisma Flow chart

al. (2020) study (the only non-interventional study included in this review), which is of moderate quality [26]. The variations in the scoring are mainly linked to limited control of confounding, absence of blinding, and minimal analytical depth. Despite some variability, all included studies met the quality criteria for inclusion in the synthesis. A detailed breakdown of the quality assessment scoring using the Checklist for Assessing the Quality of Quantitative Studies is provided in the Supplementary File 2.

Thematic Synthesis

According to the thematic synthesis of the findings from the included studies, as discussed under the subsequent heading, mucosal immune markers, in particular secretory immunoglobulin A (sIgA), significantly influence the susceptibility of both human and animal models to colonization and infection by antimicrobial-resistant (AMR) pathogens.

Mucosal sIgA as a Protective Marker Against AMR Pathogens

Increased mucosal sIgA levels were typically linked to decreased bacterial colonization, less severe illness, or improved clearance of AMR pathogens in both observational and interventional studies. For example, in a clinical human study, Clarkson et al. (2020) reported that systemic IgG responses were not predictive of illness outcome, whereas elevated baseline serum and fecal LPS-specific IgA titers were associated with protection against Shigella sonnei infection [32]. Similarly, fecal sIgA responses after oral Shigella immunization peaked after 10 days and contributed to mucosal bactericidal activity, according to Shrivastava et al. (2023) [33].

Moreover, in animal models, mucosal sIgA induction, either through probiotics or vaccination, was associated with decreased colonization by Acinetobacter baumannii and Campylobacter jejuni according to Wang et al. (2017) and Perruzza et al. (2020) studies [23, 24]. These results highlight the critical function of sIgA in blocking resistant bacteria's adherence to and invasion of the mucosa.

Mucosal Site Variation and Study Design

Most of the studies assessed gastrointestinal sIgA responses, particularly about enteric pathogens such as Shigella, Salmonella, and E. coli. For instance, Xiangchen et al. (2024) discovered that mice exposed to Lactobacillus strains had higher intestinal sIgA levels [30], which in turn reduced colonization by E. coli strains that produce extended-spectrum β-lactamase (ESBL). On the other hand, Li et al. (2020) study showed that aerosol immunization improved sIgA in the respiratory mucosa, protecting rats against strains of Acinetobacter that are resistant to multidrug [27].

Notably, interventional studies such as that of Haldar et al., (2024) and Wang et al., (2017) provided more robust causal evidence linking sIgA to AMR mitigation [23, 28], while observational study by Gudaryan et al., (2020) primarily identified correlational associations between baseline sIgA levels and bacterial infection [26].

Effects of Other Mucosal Immune Markers

Although the main focus was on sIgA, a number of studies also reported on other mucosal markers, including pIgR, TNF-α, IL-6, and IgG, providing an additional perspective of mucosal immunity. For instance, Gorain et al. (2021) discovered that resistance to multidrug-resistant E. coli was linked to the co-expression of IgA and IL-6 in the gut, indicating that cytokine-mediated immune regulation could join forces with sIgA [25].

Comparative Insights from Human and Animal Studies

While real-world evidence of mucosal immune responses and their correlation with disease and colonization resistance was presented in human studies, molecular insights and experimental control (e.g., by vaccination, microbiota manipulation) were investigated in animal models. The functional significance of mucosal sIgA in mediating protection in natural infection settings is supported by Clarkson et al. (2020), who found that individuals with higher baseline fecal and serum sIgA levels targeting Shigella lipopolysaccharide had lower incidence and severity of disease in a controlled human challenge study [32]. Similarly, after oral vaccination with live attenuated Shigella strains, mucosal sIgA responses peaked after 10 days and were linked to increased bactericidal activity against resistant strains in stool samples, according to Shrivastava et al. (2023) [33]. These suggest that, even in real-world or semi-controlled settings, mucosal immune activation, in particular sIgA production, can function as a modulatory factor and a predictive marker in host defense against AMR pathogens.

On the other hand, dietary or pharmacological interventions that altered sIgA secretion were associated with quantifiable decreases in pathogen load and inflammation in controlled animal experiments, as reported by Li et al. (2024), underscoring mucosal immune markers as modifiable predictors of AMR outcomes [31].

Ultimately, findings from these studies support the hypothesis that variations in sIgA levels and activity at mucosal surfaces significantly influence host susceptibility to antimicrobial-resistant pathogens. While sIgA emerged as a consistently relevant marker, its protective role was often potentiated by other immune effectors such as IgG and cytokines. Differences across species, mucosal sites, and study designs suggest that mucosal immune dynamics are highly context-dependent, and future research should prioritize longitudinal human studies and integrative immune profiling to better define correlates of protection.

Discussion

Confirmed Findings

This review demonstrates that mucosal secretory IgA (sIgA) consistently reduces colonization and disease caused by antimicrobial-resistant (AMR) pathogens. Across both human and animal studies, higher sIgA levels were associated with decreased bacterial adherence, reduced symptom severity, and improved pathogen clearance. Interventional studies provided particularly strong evidence: for example, Wang et al. (2017) and Perruzza et al. (2020) reported 2–5-fold increases in sIgA levels with corresponding reductions in MRSA and Campylobacter colonization [23, 24]. These findings support the biological role of sIgA as a first-line defense mechanism that excludes pathogens at mucosal surfaces before systemic infection develops [34–36].

Nuances and Complexities

The protective role of sIgA rarely acts in isolation. Several studies reported congruent elevations in cytokines such as IL-6 and TNF-α alongside sIgA, suggesting synergistic immune pathways that reinforce epithelial integrity [25, 28]. The microbiota also emerged as an important modulator: commensal organisms stimulate sIgA secretion through Th17 and polymeric IgR pathways, enhancing colonization resistance [39–41]. Restoration of microbiota in antibiotic-treated mice further demonstrated improved sIgA-mediated protection [42–44]. Moreover, site-specific effects were noted. While gastrointestinal studies consistently demonstrated protective sIgA activity,

respiratory and genitourinary models showed variable responses [27, 31], underlining that mucosal immune dynamics are highly context-dependent.

Limitations and Methodological Considerations

Despite these promising findings, several limitations must be acknowledged. Most animal studies lacked blinding or randomization, introducing risk of bias, while human trials were often underpowered and of short duration [16, 26]. Importantly, heterogeneity in study designs, mucosal sites, and outcome definitions affected the synthesis of findings. For example, inconsistent definitions of "colonization" versus "infection," varied laboratory methods for measuring sIgA, and differences in population characteristics limited comparability across studies. This variability not only complicates interpretation but also raises the possibility of selective reporting or measurement bias. Addressing such heterogeneity will be critical for strengthening future evidence [17, 18].

Implications and Future Directions

The findings of this review reinforce the translational value of targeting mucosal immunity in AMR prevention. Vaccines such as FimH and ExPEC4V, as well as sublingual formulations like Uromune (MV140), demonstrate the feasibility of mucosal immunization for reducing recurrent infections and antibiotic use [45, 46]. Probiotic interventions and microbiota-modulating strategies also offer accessible, non-antibiotic approaches to enhancing sIgA responses [10, 11, 14]. Future research should prioritize standardized measurement of mucosal markers, longitudinal human trials, and integrative immune profiling that combines sIgA with cytokines, IgG, and microbiome analysis. For clinicians, these findings emphasize the need to consider mucosal immunity in patient risk assessment and treatment planning. For researchers, they identify sIgA as a measurable endpoint for intervention trials. For policymakers, they highlight the strategic importance of investing in mucosal immunization programs as part of a comprehensive response to the AMR crisis [1–5].

Strengths of the Study

The included studies were conducted across varied geographical and clinical contexts, encompassing both human populations and animal models, which enhances the external validity of the findings. The inclusion of both observational and interventional designs allows for a well-rounded synthesis of real-world and mechanistic evidence. Most studies assessed secretory IgA quantitatively using standardized immunoassays, and several explored co-regulation with other immune markers, offering deeper insight into mucosal immune dynamics. Furthermore, the comprehensive search strategy across multiple databases and the use of an established quality assessment tool ensured methodological rigor.

Limitations of the Study

The heterogeneity in study designs, mucosal sites, and outcome definitions also affected the synthesis of findings. For example, differences in how colonization versus infection outcomes were measured, or in the techniques used to quantify sIgA, limited direct comparability. This variability increases the risk of interpretive bias and highlights the need for standardized definitions and protocols in future research.

Conclusion

This systematic review affirms that secretory IgA (sIgA) is a critical first-line defense at mucosal surfaces, where it prevents colonization and infection by antimicrobial-resistant (AMR) bacteria. Evidence from both human and animal studies consistently shows that elevated sIgA levels are associated with reduced AMR pathogen burden, with effects further strengthened by interactions with cytokines, IgG, and the microbiota.

Despite these promising findings, methodological heterogeneity—ranging from differences in study design and outcome definitions to variations in sIgA measurement—limits comparability and underscores the need for standardized protocols in future work. Addressing these limitations will be critical for translating sIgA from a biological marker into a reliable predictor of protection against AMR.

The implications of this review extend across stakeholders. For clinicians, sIgA offers a potential biomarker for risk stratification in patients vulnerable to resistant infections. For researchers, it highlights the need for longitudinal human studies that integrate immune profiling with microbiome analysis. For policymakers, the findings emphasize the value of investing in mucosal vaccine development, probiotic interventions, and other preventive strategies as part of global AMR control efforts.

Ultimately, strength------ening mucosal immunity—particularly through interventions that boost sIgA—represents a promising, non-antibiotic pathway for reducing the global burden of antimicrobial resistance.

Supplementary materials

The Supplementary information includes tables:

• Supplementary Table 1. Charateristics of the included studies

• Supplementary Table 2. Quality Assessment.

This supplemental material has been provided by the authors to give readers additional information about their work.

The file can be accessed using the following links:

• https://www.editorialpark.com/download/article-supp/734/Supplement-File-1-(Charateristics-of-the-included-studies)-(1).xlsx;

• https://www.editorialpark.com/download/article-supp/735/Supplement-File-2-(Quality-Assessment)-(1).xlsx.

Author Contributions: Conceptualization, O.P.L. and A.O.O.; methodology, O.P.L., A.O.O., U.C.T. and C. R.; validation, M.J.B., M.O.A. and U.M.O.; formal analysis, O.P.L.,A.O.O. and I.O.O.; investigation,O.P.L., U.C.T., and M.J.B.; resources, M.J.B. and I.O.O.; data curation, C.R. and M. O. A.; writing – original draft preparation,O.P.L., A.O.O., U.C.T. and C. R.; writing – review and editing, O.P.L., U.C.T., M.O.A.; visualization, C.R. and U.M.O.; supervision, O.P.L. and I. O.O.; project administration, O.P.L.; funding acquisition, O.P.L. All authors have read and agreed to the published version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: The authors acknowledge all others who played a beneficial role in completing this manuscript.

Funding: None.

Data availability statement: No datasets were generated or analysed during the current study.

Artificial Intelligence (AI) Disclosure Statement: AI-Unassisted Work.

References

1. World Health Organization. *Antimicrobial resistance.* 2023 Nov 21. Available from: https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance
2. Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, Han C, Bisignano C, Rao P, Wool E, Johnson SC, Browne AJ, Chipeta MG, Fell F, Hackett S, Hanefeld J, Etmon M, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet*. 2022;399(10325):629–655. https://doi.org/10.1016/S0140-6736(21)02724-0.
3. Ahmed SK, Hussein S, Qurbani K, Ibrahim RH, Fareeq A, Mahmood KA, Mohamed MG. Antimicrobial resistance: Impacts, challenges, and future prospects. *J Med Surg Public Health.* 2024;2:100081. https://doi.org/10.1016/j.glmedi.2024.100081
4. Salam MA, Al Amin MY, Salam MT, Pawar JS, Akhter N, Rabaan AA, Alqumber MAA. Antimicrobial resistance: A growing serious threat to global public health. Healthcare. 2023;11:1946. https://doi.org/10.3390/healthcare11131946
5. Jacobs A. The global threat of antibiotic resistance. *The New York Times.* 2024 Sep 26. Available from: https://www.nytimes.com/2024/09/26/health/united-nations-drug-resistance-antibiotics.html
6. Nyathi N, Ndhlovu D, Rabvukwa E, Ndudzo A. Improving antimicrobial stewardship in human health using probiotics and their derivatives. In: Probiotics and Their Derivatives – Applications in Human and Animal Health. IntechOpen; 2024. https://doi.org/10.5772/intechopen.112497
7. Corthésy B. Role of secretory immunoglobulin A and secretory component in the protection of mucosal surfaces. Future Microbiol. 2010;5(5):817–829. https://doi.org/10.2217/fmb.10.39
8. Zhou X, Wu Y, Zhu Z, Lu C, Zhang C, Zeng L, Xie F, Zhang L, Zhou F. Mucosal immune response in biology, disease prevention and treatment. *Signal Transduct Target Ther.* 2025;10(1):7. https://doi.org/10.1038/s41392-024-02043-4
9. Oluwasola A, Otegbayo J, Ola S, Ebili H, Afolabi A, Odaibo G. Correlation of serum anti-helicobacter pylori immunoglobulin A (IgA) with histological parameters of chronic gastritis in Ibadan, Nigeria. *Ann Ib Postgrad Med.* 2012;10(1):18–24.
10. Kusumo PD, Bela B, Wibowo H, Munasir Z, Surono IS. Lactobacillus plantarum IS-10506 supplementation increases faecal sIgA and immune response in children younger than two years. *Benef Microbes.* 2019;10(3):245–252. https://doi.org/10.3920/BM2017.0178
11. Ku S, Haque MA, Jang MJ, Ahn J, Choe D, Jeon JI, Park MS. The role of Bifidobacterium in longevity and the future of probiotics. *Food Sci Biotechnol.* 2024;33(9):2097–2110. https://doi.org/10.1007/s10068-024-01631-y
12. Biagi E, Franceschi C, Rampelli S, Capri M, Brigidi P, Candela M. Gut microbiota and extreme longevity. *Curr Biol*. 2016;26(11):1480–1485. https://doi.org/10.1016/j.cub.2016.04.016
13. Rampelli S, Soverini M, D'Amico F, Barone M, Tavella T, Monti D, Capri M, Astolfi A, Brigidi P, Biagi E, Franceschi C, Turroni S, Candela M. Shotgun metagenomics of gut microbiota in humans with up to extreme longevity and the increasing role of xenobiotic degradation. mSystems. 2020;5(2):e00124 20. https://doi.org/10.1128/mSystems.00124 20
14. Chen Chi Chung, Lai Chih Cheng, Huang Hui Ling, Huang Wen Yu, Toh Han Siong, Weng Tzu Chieh, Chuang Yin Ching, Lu Ying Chen, Tang Hung Jen. Antimicrobial activity of Lactobacillus species against carbapenem resistant Enterobacteriaceae. *Front Microbiol.* 2019;10:789. https://doi.org/10.3389/fmicb.2019.00789
15. Yan Rushuang, Lu Ye, Wu Xiaoqing, Yu Peihao, Lan Peng, Wu Xueqing, Jiang Yan, Li Qi, Pi Xionge, Liu Wei, Zhou Jiancang, Yu Yunsong. Anticolonization of carbapenem resistant Klebsiella pneumoniae by Lactobacillus plantarum LP1812 through accumulated acetic acid in mice intestinal. *Front Cell Infect Microbiol.* 2021;11:804253. https://doi.org/10.3389/fcimb.2021.804253
16. Blagojevic Christina, Brown Kevin A, Diong Christina, Fridman Daniel J, Johnstone Jennie, Langford Bradley J, Lee Samantha M, MacFadden Derek R, Schwartz Kevin L, Daneman Nick. Long-term risk of infection among patients colonized with antimicrobial-resistant pathogens: a population-wide cohort study. Open Forum Infect Dis. 2024;11(12):ofae712. https://doi.org/10.1093/ofid/ofae712
17. Virgo M, Mostowy S, Ho BT. Emerging models to study competitive interactions within bacterial communities. *Trends Microbiol*. 2025;33(6):688–700. https://doi.org/10.1016/j.tim.2024.12.009
18. Tao L, Reese TA. Making mouse models that reflect human immune responses. *Trends Immunol.* 2017;38(3):181–193. https://doi.org/10.1016/j.it.2016.12.007
19. Page Matthew J, McKenzie Joanne E, Bossuyt Patrick M, Boutron Isabelle, Hoffmann Tammy C, Mulrow Cynthia D, Shamseer Larissa, Tetzlaff Jennifer M, Akl Elie A, Brennan Sue E, Chou Roger, Glanville Julie, Grimshaw Jeremy M, Hróbjartsson Asbjørn, Lalu Manoj M, Li Tianjing, Loder Elizabeth W, Mayo Wilson Evan, McDonald Steve, McGuinness Luke A, Stewart Lesley A, Thomas James, Tricco Andrea C, Welch Vivian A, Whiting Penny, Moher David. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. https://doi.org/10.1136/bmj.n71
20. Hosseini MS, Jahanshahlou F, Akbarzadeh MA, Zarei M, Vaez-Gharamaleki Y. Formulating research questions for evidence-based studies. J Med Surg Public Health. 2024;2:100046. https://doi.org/10.1016/j.glmedi.2023.100046
21. Kmet LM, Lee RC, Cook LS. Standard quality assessment criteria for evaluating primary research papers from a variety of fields. Edmonton: Alberta Heritage Foundation for Medical Research; 2004. https://doi.org/10.7939/R37M04F16
22. Silvani MI, Werder R, Perret C. The influence of blue light on sleep, performance and wellbeing in young adults: A systematic review. Front Physiol. 2022;13:943108. https://doi.org/10.3389/fphys.2022.943108
23. Wang Maopeng, Gong Shengjie, Du Shouwen, Zhu Yilong, Rong Fengjun, Pan Rongrong, Di Yang, Li Chang, Ren Dayong, Jin Ningyi. The effect of immunoregulation of Streptococcus lactis L16 strain upon Staphylococcus aureus infection. *BMC Microbiol.* 2017;17(1):130. https://doi.org/10.1186/s12866-017-1038-y
24. Perruzza L, Jaconi S, Lombardo G, Pinna D, Strati F, Morone D, Seehusen F, Hu Y, Bajoria S, Xiong J, Kumru OS, Joshi SB, Volkin DB, Piantanida R, Benigni F, Grassi F, Corti D, Pizzuto MS. Prophylactic Activity of Orally Administered FliD-Reactive Monoclonal SIgA Against Campylobacter Infection. *Front Immunol.* 2020;11:1011. https://doi.org/10.3389/fimmu.2020.01011
25. Gorain C, Khan A, Singh A, Mondal S, Mallick AI. Bioengineering of LAB vector expressing Haemolysin co-regulated protein (Hcp): a strategic approach to control gut colonization of Campylobacter jejuni in a murine model. *Gut Pathog.* 2021;13(1):48. https://doi.org/10.1186/s13099-021-00444-2
26. Gudaryan AA, Maschenko IS, Dorogina AS, Shyrinkin SV. Approaches in correction of some changes of the immunologic disorders in patients with rapidly progressing generalized periodontitis with different drug susceptibility of the periodontal bacteria. Medicni

Perspektivi. 2020;25(4):138–146. https://doi.org/10.26641/2307-0404.2020.4.221409
27. Li S, Chen DQ, Ji L, Sun S, Jin Z, Jin ZL, Sun HW, Zeng H, Zhang WJ, Lu DS, Luo P, Zhao AN, Luo J, Zou QM, Li HB. Development of Different Methods for Preparing Acinetobacter baumannii Outer Membrane Vesicles Vaccine: Impact of Preparation Method on Protective Efficacy. *Front Immunol.* 2020;11:1069. https://doi.org/10.3389/fimmu.2020.01069
28. Haldar R, Dhar A, Ganguli D, Chakraborty S, Pal A, Banik G, Miyoshi SI, Das S. A candidate glycoconjugate vaccine induces protective antibodies in the serum and intestinal secretions, antibody recall response and memory T cells and protects against both typhoidal and non-typhoidal Salmonella serovars. *Front Immunol.* 2024;14:1304170. https://doi.org/10.3389/fimmu.2023.1304170
29. Zhang F, Bai J, Zheng Y, Liang S, Lei L, Deng X, Li W, Liu P, Yang G, Ren Y. Investigation of the Optimum Preparation of Peach Gum Polysaccharides and the In Vivo and In Vitro Therapeutic Effects on Acute Pyelonephritis. *Evid Based Complement Alternat Med.* 2019;2019:2729343. https://doi.org/10.1155/2019/2729343
30. Xiangchen GU, Meisi Q, Lin X, Min C, Yueyi D, Changming Z, Guihua J, Chen W, Yi W. Individualized Traditional Chinese Medicine treatment antibiotics for recurrent urinary tract infections: a multicenter, randomized controlled study. *J Tradit Chin Med*. 2024;44(3):524–529. https://doi.org/10.19852/j.cnki.jtcm.20231024.001
31. Li Z, Wang X, Zheng D, Han F, Li Y, Zhou H, Li J, Cui W, Jiang Y, Wang X, Xie W, Tang L. Oral-delivery Lactococcus lactis expressing cherry fusion lactoferrin peptides against infection of avian pathogenic Escherichia coli in chickens. *Poult Sci.* 2025;104(1):104637. https://doi.org/10.1016/j.psj.2024.104637
32. Clarkson KA, Frenck RW Jr, Dickey M, Suvarnapunya AE, Chandrasekaran L, Weerts HP, Heaney CD, McNeal M, Detizio K, Parker S, Hoeper A, Bourgeois AL, Porter CK, Venkatesan MM, Kaminski RW. Immune Response Characterization after Controlled Infection with Lyophilized Shigella sonnei 53G. mSphere. 2020;5(5):e00988-19. https://doi.org/10.1128/mSphere.00988-19
33. Shrivastava S, Agnememel AB, Ndungo E, Islam D, Liang Y, Frenck RW, Pasetti MF. Oral immunization with Shigella sonnei WRSs2 and WRSs3 vaccine strains elicits systemic and mucosal antibodies with functional anti-microbial activity. mSphere. 2024;9(3):e00419-23. https://doi.org/10.1128/msphere.00419-23
34. Strugnell R, Wijburg O. The role of secretory antibodies in infection immunity. *Nat Rev Microbiol.* 2010;8(8):656–667. https://doi.org/10.1038/nrmicro2384
35. Johansson MEV, Phillipson M, Petersson J, Velcich A, Holm L, Hansson GC. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci U S A*. 2008;105(39):15064–15069. https://doi.org/10.1073/pnas.0803124105
36. Tokuhara Daisuke, Yuki Yoshikazu, Nochi Tomonori, Kodama Toshio, Mejima Mio, Kurokawa Shiho, Takahashi Yuko, Nanno Masanobu, Nakanishi Ushio, Takaiwa Fumio, Honda Takeshi, Kiyono Hiroshi. Secretory IgA-mediated protection against Vibrio cholerae and heat-labile enterotoxin-producing enterotoxigenic Escherichia coli by rice-based vaccine. *Proc Natl Acad Sci U S A*. 2010;107(19):8794–8799. https://doi.org/10.1073/pnas.0914121107
37. Bunker Jeffrey J, Erickson Steven A, Flynn Theodore M, Henry Carole, Koval Jason C, Meisel Marlies, Jabri Bana, Antonopoulos Dionysios A, Wilson Patrick C, Bendelac Albert. Natural polyreactive IgA antibodies coat the intestinal microbiota. Science. 2017;358(6361):eaan6619. https://doi.org/10.1126/science.aan6619
38. van Dalen R, Elsherbini AMA, Harms M, Perret M, Zaat SAJ, de Steenhuijsen Piters WAA. Secretory IgA impacts the microbiota density in the human nose. *Microbiome*. 2023;11:233. https://doi.org/10.1186/s40168-023-01675-y
39. Ivanov Ivaylo I, Atarashi Koji, Manel Nicolas, Brodie Eoin L, Shima Tatsuichiro, Karaoz Ulas, Wei Dongguang, Goldfarb Katherine C, Santee Clark A, Lynch Susan V, Tanoue Takeshi, Imaoka Akemi, Itoh Kikuji, Takeda Kiyoshi, Umesaki Yoshinori, Honda Kenya, Littman Dan R. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*. 2009;139(3):485–498. https://doi.org/10.1016/j.cell.2009.09.033
40. Atarashi Koji, Tanoue Takeshi, Oshima Kenshiro, Suda Wataru, Nagano Yuji, Nishikawa Hiroyoshi, Fukuda Shinji, Saito Takuro, Narushima Seiko, Hase Koji, Kim Sangwan, Fritz Joëlle V, Wilmes Paul, Ueha Satoshi, Matsushima Kouji, Ohno Hiroshi, Olle Bernat, Sakaguchi Shimon, Taniguchi Tadatsugu, Morita Hidetoshi, Hattori Masahira, Honda Kenya. Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature*. 2013;500(7461):232–236. https://doi.org/10.1038/nature12331
41. Palm Noel W, de Zoete Mark R, Cullen Travis W, Barry Nicholas A, Stefanowski Jared, Hao Lili, Degnan Patrick H, Hu Ji, Peter Irene, Zhang Wen. Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell*. 2014;158(5):1000–1010. https://doi.org/10.1016/j.cell.2014.08.006
42. Pietrzak B, Tomela K, Olejnik-Schmidt A, Mackiewicz A, Schmidt M. Secretory IgA in intestinal mucosal secretions as an adaptive barrier against microbial cells. Int J Mol Sci. 2020;21(23):9254. https://doi.org/10.3390/ijms21239254
43. Hodgkinson Alison J, Cakebread Julie, Callaghan Megan, Harris Paul, Brunt Rachel, Anderson Rachel C, Armstrong Kelly M, Haigh Brendan. Comparative innate immune interactions of human and bovine secretory IgA with pathogenic and non pathogenic bacteria. *Dev Comp Immunol.* 2017;68:21–25. https://doi.org/10.1016/j.dci.2016.11.012
44. Gorrell Rebecca J, Wijburg Odilia L C, Pedersen John S, Walduck Anna K, Kwok Terry, Strugnell Richard A, Robins Browne Roy M. Contribution of secretory antibodies to intestinal mucosal immunity against Helicobacter pylori. *Infect Immun.* 2013;81(10):3880–3893. https://doi.org/10.1128/IAI.01424-12
45. Eldridge Gary R, Hughey Heidi, Rosenberger Lois, Martin Steven M, Shapiro Andrew M, D'Antonio Elizabeth, Krejci Kent G, Shore Neal, Peterson James, Lukes Andrea S, Starks Courtney M. Safety and immunogenicity of an adjuvanted Escherichia coli adhesin vaccine in healthy women with and without histories of recurrent urinary tract infections: results from a first in human phase 1 study. *Hum Vaccin Immunother.* 2021;17(5):1262–1270.https://doi.org/10.1080/21645515.2020.1834807
46. Nickel JC, Kelly K-L, Griffin A, Elterman DS, Clark-Pereira J, Doiron RC. MV140 sublingual vaccine reduces recurrent urinary tract infection in women: Results from the first North American clinical experience study. *Can Urol Assoc J.* 2023;18(2):25–31. https://doi.org/10.5489/cuaj.8455

JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Review Article DOI: https://doi.org/10.23950/jcmk/17533

Extracellular Vesicles and Lyophilization: Getting Functionally Stable Cell-free Biomedical Products

Abay Baigenzhin[1], Elmira Chuvakova[2], Aigerim Zhakupova[3], Aizhan Akhayeva[3], Anastassiya Ganina[3]

[1]Chairman of the Board, JSC "National Scientific Medical Center", Astana, Kazakhstan
[2]Deputy of Science, JSC "National Scientific Medical Center", Astana, Kazakhstan
[3]Department of Cell Technologies and Transplantation, JSC "National Scientific Medical Center", Astana, Kazakhstan

Received: 2025-11-27.
Accepted: 2025-12-11.

J Clin Med Kaz 2025; 22(6): 117–125

Corresponding author:
Aigerim Zhakupova
E-mail: s.aigerim.87@mail.ru.
ORCID: 0009-0007-2042-0036.

Abstract

During the last decade, clinical application of extracellular vesicles (EVs) is of growing interest. Despite the progress in exploring the therapeutic potential of EVs, e.g. as disease markers or the carriers for therapeutic substances, it is important to identify proper storage conditions – this issue is indeed challenging. A subtype of EVs known as exosomes is of great importance in the therapeutic applications because they participate in the regulation of intercellular communication. Currently, exosomes are considered as a promising tool in the regenerative medicine. The therapeutic potential of exosomes and other subtypes of EVs, especially for their using in immunomodulation and drug delivery, dictates great attention to the methods for their storage, in particular for a long periods of time. Lyophilization is one of the best such methods designed to preserve cell-free EV-based products. In our mini-review, we discuss the main methods developed for stabilizing cell-free products and getting stable solid forms of EVs that are capable for long-term storage. We also point out that the methods need to be following both the ease of transportation and the retaining the functionally important properties for in vivo applications. The development of optimal protocols for storing the EVs are therefore crucially important for warranting that structural and functional integrity of EVs, exosomes in particular, are maintained intact or at least modified as less as possible. For comprehensiveness of our review, we refer to original studies which investigated how the storage temperature and freezing methods may affect stability of the final EV product. We summarize advances in the area of freeze-drying EVs (exosomes), the selection of optimal process parameters and lyoprotectants, the interplay between lyophilization parameters and specific functional properties of exosomes, and the preservation of their biological activity after reconstitution before application in vivo.

Keywords: extracellular vesicles, exosomes, lyophilization, cryopreservation, biomedical cell-free products

1. Introduction

Cells are constantly engaged in the exchange of molecular signals, and one of the most efficient ways they communicate is through the release of extracellular vesicles (EVs). This form of intercellular interaction relies on a multistep intracellular machinery that governs the formation, packaging, and secretion of vesicular structures [1]. At the final stages of their maturation, EVs appear as lipid-bounded particles that transport a broad spectrum of biologically active components. It includes proteins, diverse RNA species and their fragments, nucleic acids, metabolites, and lipids synthesized inside the parent cell [2]. The complexity and heterogeneity of the cargo is considered

a key feature of EVs because it extends their functional versatility and allows them to regulate a wide array of biological processes. Indeed, they are important mediators in the cell-to-cell inter-talk because they transport various ready-to-use proteins (including receptors and ion channels). In turn, it makes possible to perform different functions by the same carriage. For example, EVs contain certain types of RNA (mRNA, miRNA, siRNA, etc.) and can also contain DNA fragments that is critical in inter-cellular signaling (paracrine regulation). In addition, they can affect function of distant cells at the systemic level via circulatory system. Under physiological conditions, EVs contribute to tissue homeostasis and support normal regulatory pathways. However, their involvement becomes even more pronounced during pathological events, when cells release vesicles enriched with stress-associated or damage-related molecules. In this scenario, the function of recipient cells can be regulated both directly (via fusion to the cell membrane and releasing the content inside the cell) and indirectly (via activation of extracellular signaling molecules).

Virtually all cell types are capable of releasing EVs, making this mechanism one of the most energy-efficient routes of cell-to-cell signaling aside from direct electrical communication. Nevertheless, EVs generated by terminally differentiated cells tend to carry limited regenerative potential. In contrast, vesicles secreted by stem cells have been shown to reproduce many of the therapeutic functions attributed to mesenchymal stem cells (MSCs) themselves [2, 3]. Current classifications describe three main categories of EVs – microvesicles, exosomes, and apoptotic bodies – which differ in their origin, release pathways, morphology, cargo composition, and functional features. The main characteristics and differences between the different types of EVs, including their functional role in intercellular signaling, are summarized in Table 1.

Among different EV types, exosomes represent the most mobile and biologically active population. They are widely detectable in blood and can also be isolated from various body fluids such as urine, breast milk, saliva, bile, lymph, and cerebrospinal fluid. The molecular profile of exosomes includes numerous signaling factors that demonstrate selective tropism toward particular cell types, with their composition strongly influenced by the microenvironment of the secreting cells [1, 2, 3]. The simultaneous circulation of diverse exosome populations in extracellular fluid allows fine modulation of multiple recipient cells (and diverse types of cells) at once. When cells experience stress or damage, they release exosomes enriched with pathological molecular signatures, including proteins and miRNAs characteristic of disease states. This explains their pivotal contribution to tumor progression and metastasis but has also been proven as a key player in numerous non-cancerous disorders [4, 5]. Consequently, exosomes have emerged as promising diagnostic and prognostic biomarkers, with high relevance for personalized therapeutic decision-making.

Table 1 Main characteristics of intercellular vesicles of various types involved in intercellular signaling

Properties	Exosomes	Microvesicular bodies	Apoptotic cells
Size	30-150 nm	100-1000 nm	500-2000 nm
Origin	Endosomal pathway (exocytosis of microvesicular bodies)	Detachment from plasmatic membrane	Cell fragmentation in apoptosis
Composition	Proteins, lipids, RNA, microRNA	Membrane-bound proteins, lipids, RNA	Cytoplasmic fragments, organelles, DNA
Functions	Intercellular communication, immune regulation	Transfer cell signaling	Removal of cell debris
Protein markers	CD9, CD63, CD81, Alix, TSG101, Flotillin 1,2	Matrix metalloproteases, ARF6, CD40	Phosphati-dylserine, calreticulin, CD45, Caspase 3, C1q
The mechanism of release	Exocytosis	Budding from the cell	Cell damage and destruction
Nature	Endogenic	Endogenic	Endogenic

Exosomes also interact with immune cells by presenting antigen structures on their surface and can influence recipient cells directly after fusion with their plasma membranes. These properties open possibilities for engineering therapeutic exosomes loaded with tailored protein or nucleic-acid cargo. The exosomes preloaded by biologically active molecules can be targeted to specific tissues to support their protection and repair [1, 2]. Their involvement in neural regeneration – such as myelin formation, neuronal outgrowth, and the recovery of injured glial and neuronal cells – has been demonstrated in several studies with in vitro models [6, 7]. Recent studies suggest their active contribution to the regenerative capacity of mesenchymal stem cells. For example, the recovery of myocardial tissue through the reparation of damaged cardiomyocytes or triggering differentiation of immature cells to cardiomyocytes [8] or the recovery of retina in patients with diabetes mellitus [9] can be mediated by exosomal transport.

The high potential for using EVs for diagnostic and clinical purposes, along with their direct involvement into the intercellular communication, signaling, and molecular transfer, opens the possibility of using them in biomedicine [10]. EVs are currently gaining increasing attention as cell-free platforms capable of the delivery of drugs, cellular therapeutics, and clinical biomarkers [1, 2, 11]. Extracellular vesicles are very stable in vivo and they are typically not rejected when crossing biological barriers. The loading capacity of the exosomes allows for delivery of genetic information between cells, which enables sharing of epigenetic changes and affecting to recipient cells by genetic reprogramming [7, 11]. However, there are some limiting factors that should be considered in the context of therapeutic applications of EVs. To be used in clinical settings, EVs must be produced in sufficient quantities while maintaining their viability intact or modified in minor extent. Large-scale production of EVs while maintaining structural and functional integrity remains challenging. One of the important problems is that their long-term preservation requires specific storage techniques to prevent degradation of the carrier itself or cargo loss in time [4].

Published studies indicate that EVs behave similarly to living cells regarding sensitivity of their membrane structure (protein- and lipid-based content, morphology, etc.) to environmental conditions. Variations in storage temperatures have been shown to affect both vesicle integrity, functional recovery after thawing, fusion properties and other key features of the EVs [3, 10, 11]. Once the EVs are produced, the storage temperature significantly affects their recovery and integrity, showing that different temperature ranges promote different alterations in the EVs [5, 12]. Temperature-dependent structural changes highlight the need for proper preservation methods, among which two strategies dominate in current practice. First are deep-freezing

techniques such as cryopreservation, the second are dehydration techniques such as freeze-drying (lyophilization). In our mini-review, both approaches are examined in detail, with emphasis on their advantages, constraints, and practical implications for the long-term storage of extracellular vesicles.

2. Cryopreservation

Cryopreservation has become one of the most widely applied strategies for maintaining biological materials in a functional state over extended periods. Its routine use in laboratory and clinical practice allows long-term storage of diverse biospecimens while retaining their viability and physiological activity. The availability of cryopreserved samples on demand – including cells, tissues, and complex preparations accumulated in biobanks – significantly simplifies research workflow and ensures constant access to standardized material for analytical and transplantation purposes. This eliminates the necessity for frequent procurement of freshly isolated samples and enables thorough quality assessment prior to their clinical use. Stem cells, for instance, can be stored for prolonged periods without a measurable reduction in their regenerative or therapeutic properties, which is crucial for regenerative medicine and anticancer applications. Likewise, cryogenic storage of embryos and oocytes has long been an essential component of reproductive technologies [1, 4, 13, 14]. In addition, the proven feasibility of long-term storage of microbial cultures, established cell lines, and tissue samples provides a stable resource for research activity and supports continuous advancement in biomedicine. Preservation of genetic material of rare microbial strains plays an equally important role, as it ensures lineage continuity and creates opportunities for further genetic engineering applicable in corresponding research fields [2, 11, 15].

Cryopreservation can theoretically be applied to nearly any biological substrate — from cell-free and single-cell preparations to tissues, organs, and bacterial cultures. However, because biological material consists predominantly of water, exposure to subzero temperatures inevitably triggers ice crystallization. The formation of intracellular and extracellular crystals is one of the major factors contributing to irreversible damage, compromising membrane integrity and internal organization of cells and vesicles. Repeated freeze-thaw cycles further exacerbate this effect by reducing the concentration of bioactive molecules. In addition, these cycles induce degradation of RNA and proteins, and promote aggregation of the cellular content hindering its functional capacity. These forms of cryo-injury highlight the necessity of using cryoprotective compounds capable of mitigating structural and biochemical damage during freezing [6, 11, 16].

Cryoprotectants allow for avoiding physical effects (crystallization), osmotic damage, chemical changes like variation in pH and cytoplasmic ion content. They also maintain protein molecular stability and the spatial organization of intracellular membranes (if any cell organelles are present in a product) and outer membrane [1, 3, 17]. Two types of cryoprotectants are used now: membrane-permeable and membrane-impermeable. Membrane-permeable cryoprotectants, such as dimethyl sulfoxide (DMSO), are substances with relatively low molecular weight. Accordingly, this property facilitates their permeation through double-layer membranes without involvement of passive or active transmembrane transport. The cryoprotectants of this type stabilize the vesicular membrane from the inside. In contrast, non-permeant cryoprotectants, such as sucrose and trehalose, can not cross the membrane due to their high molecular weight. They therefore reside outside the EV membranes, mostly being incorporated into the glycocalix and affecting the structural properties of the cell membrane [4, 5, 18].

Despite the advantages of cryogenic preservation, several limitations persist. Reports repeatedly describe cytotoxicity and impairment of signaling pathways associated with the use of permeable cryoprotectants, particularly at higher concentrations or prolonged exposure [13, 14, 19]. Technical and logistical issues may also compromise sample integrity and stability. Interruptions in power supply, failure of refrigeration systems, or delays in refilling liquid nitrogen tanks may lead to unexpected sample loss, operational delays, and substantial financial burden [7, 20]. Furthermore, the requirement for uninterrupted storage at ultralow temperatures (for example, at -80 degrees Celsius or below) makes cryopreservation an inherently resource-dependent and costly approach.

3. Lyophilization (Freeze-drying)

Compared with cryopreservation, which is better suited for storing bulk cell preparations and tissues, lyophilization has emerged as one of the most suitable approaches for stabilizing cell-derived bioproducts such as peptides, proteins, extracellular vesicles, and vaccines [3, 21]. A major advantage of this method is that the biomaterial can be converted into a dry, solid form that remains its functional features for extended periods. Numerous studies show that when freeze-drying is performed

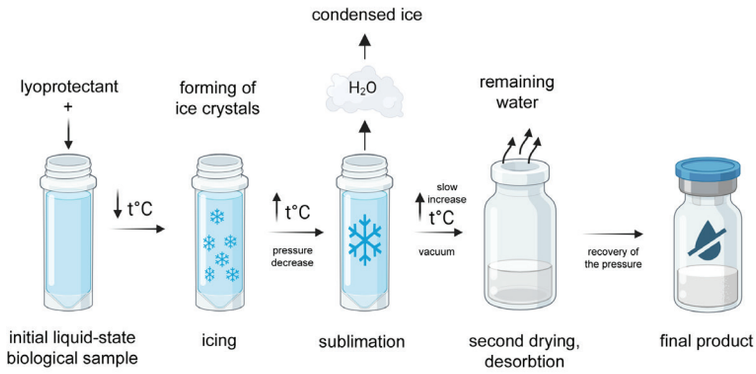


Figure 1 – The main stages of the cell lyophilization process

Image was created in Biorender.com.

in the presence of suitable cryoprotective agents, EVs preserve their morphology, size distribution, protein cargo, and overall biological activity. Importantly, lyophilized EVs maintain comparable rates of uptake and exhibit functional properties similar to freshly isolated vesicles [1, 11, 22, 23]. The principal stages of the process are illustrated in Figure 1.

First, freeze-drying is based on freezing water and then removing it away from the sample. Sublimation is the key step in the whole freeze-and-dry process, whereby water almost instantly transforms from solid state to the vapor, bypassing the liquid phase. This occurs under high pressure and at temperatures below 0°C. The process consists of three stages, which are schematically represented in Figure 1, each with different effect to the final quality of the product (see below) and different duration. During the first stage, which is the freezing, the sample solidifies with formation of ice crystals. Next, the second stage starts by the placement of the frozen sample to the vacuum allowing the ice to transform into a vapor (gas) without entering the liquid phase. The sublimation is accelerated in a heated condition. A low temperature condenser converts the evaporated water released from the vacuum chamber into a solid state [11, 21, 24, 25].

It should be noted that before freezing the biosamples must undergo specific pre-treatment. This includes adjusting the product concentration, introducing of components to enhance product stability, reducing the amount of solvent with high vapor pressure, and other technological measures. During the freezing stage, pure water ice crystals start to form and concentrate in the sample by solidification – they may present in amorphous state or in crystals or in combination of the two states [5, 10, 13, 23]. The sample temperature is lowered below freezing to achieve solidification. When temperature goes down the freeze-point of the sample's medium, the thermodynamically favorable ice clusters have higher probability to form. These clusters both act as the nuclei for initiation of medium crystallization and stimulate the crystallization process [1, 4, 5, 26]. After nucleation, ice crystals grow by increasing their volume; in a sample there are many nucleation sites appearing simultaneously, which provides nearly homogeneous freezing of the sample.

The freezing stage has direct influence on subsequent stages of the whole cycle of lyophilization. For example, higher super-cooling (lower crystal-nucleation temperature) promotes faster nucleation, resulting in appearance of larger amount of small nuclei, comprising in total much larger surface area. This prolongs primary drying time but increases the rate of desorption of unfrozen water during the next stage of lyophilization (second drying) [12, 19, 27-30]. In contrast, lower super-cooling with high crystal-nucleation temperature facilitates occurrence of much smaller amount of crystals but each crystal is larger. Altogether, it reduces the time needed for primary drying and prolongs the duration of secondary drying. However, depending on the design and capacity of the freezing apparatus, different vials of biosamples may cool under non-equal conditions. Subsequently, it may lead to variations in the rate of freezing or quality of final product between vials and batches [5, 13, 27, 31].

Of note, the cell debris, remnants of disrupted cells, and organelles should be discarded prior to the process because they do not bear functional load, need longer recovery time, and have reduced stability [14, 32]. The progression of ice formation leads to decrease in free water content. Because most of chemicals are not soluble in solid phase of the water, this in turn makes all the solutes dissolved in the liquid phase of sample's medium to be increasingly concentrated. The elevated concentrations of dissolved solutes in freezing sample contribute to the increasing in viscosity of the liquid medium until solid ice can no longer form. Because freezing conditions strongly contribute to the structure of the frozen matrix and the quality of the final product, the entire process requires permanent control, especially its first stage as most critical one [11, 21, 28, 33].

The primary drying is the main step of the whole drying process. It typically takes a longer time compared to other phases of the freeze-drying process, making it another critical step in the developing and optimizing the lyophilization process [4, 5, 22, 29]. Since primary drying step is in progress, it is needed to reduce pressure and provide sufficient heat to the sample to sublimate the ice [1, 24, 25]. During the initial drying stage, approximately 95% of the water in the material is subliming. This process can take a significant amount of time, as too much heat can alter the structure of the material.

Subsequently, while the primary drying step is occurring, the ice is continuously subliming into the vacuum. Vacuum accelerates sublimation, making it useful for targeted drying. The driving force for sublimation is the actual difference between the temperatures of ice and the surface of condenser [11, 35-38]. The heating is required for effective sublimation while the physical way it is transferred into the sample (e.g. radiation or convection) does not play a role. The heat supplied should be high enough to promote ice sublimation, but the sample temperature should remain lower enough to prevent the precipitate from collapsing after freeze-drying [3, 15, 16, 20]. The cold chamber of the condenser and its plates all represent a surface to re-solidification of water vapor. The primary role of condenser is to prevent entering the vapor to the vacuum pump, otherwise the performance of the system may be affected [18, 24, 25, 34, 39]. The chamber must be pressurized fairly below the level of saturated vapors throughout the process to promote ice sublimation. The pressure is controlled using a partial vacuum. If the pressure is lowered, it elevates the actual difference in vapor pressure between ice and condenser, but it also significantly reduces the convective heat flow, which slows down sublimation [6, 11, 19, 40-42].

The next step is secondary drying. During this stage, water is gradually evaporated by heating and vacuum [15]. However, the amount of unfrozen water may still be significant (20-50%) during this process [5, 11]. The final moisture content of the lyophilisate is a critical parameter, as it determines the product's on-shelf stability. During this stage, the temperature must be set at much higher levels compared to the settings during primary drying. This is needed for disrupting the physicochemical interactions between water molecules still adhered to the frozen sample. Because the ice is removed in primary drying, this reduces both risks of melting and sample decomposing to minimal levels [3, 24, 43, 44]. During this stage, the pressure is also typically reduced to stimulate desorption. However, the temperature increase is preferably carried out slowly (≤1 °C/min) to avoid decomposition.

One or more drying phases may be carried out at any suitable temperature. The pressure at which the drying phase is carried out may also vary because it depends on the design of the chamber, vacuumizing rate and intensity, etc. As a result, the final product is completely dry, ensuring greater stability and a longer shelf life.

The parameters of secondary drying determine the long-term stability of the final lyophilized product. The optimal residual moisture content is commonly expected to be between 0.5% and 3% [12], although certain biological materials demonstrate

better functional recovery when slightly higher moisture levels are retained [16, 22, 38]. Thus, proper optimization of secondary drying is essential for preserving structural integrity, ensuring biological activity, and achieving efficient rehydration upon use. As well, optimal residual water content is crucial for the successful recovery of lyophilized biomaterials and cells.

4. Lyoprotectants

Lyoprotectants represent a broad group of chemical compounds that safeguard biological preparations during freezing and drying by altering the physical properties of the surrounding medium. A key mechanism of their action is the ability to markedly increase viscosity, which slows or even suppresses the formation of crystalline ice. Under these conditions, water solidifies predominantly in an amorphous, non-crystalline state, forming a glass-like matrix (in contrast to typical lattice-structured ice under normal conditions). This is essential because crystal growth can deform and rupture lipid membranes, and therefore damage vesicle structure at the very initial stage of production. During sublimation, when bulk water is removed, the phospholipid bilayer tends to lose structural cohesion. Sugar-based lyoprotectants compensate for this by replacing water molecules and stabilizing the interface between the inner and outer leaflets of the membrane [4, 16, 25]. In this way, lyoprotectants prevent collapse or shrinkage of vesicles and sustain their native architecture.

A large variety of lyoprotectant formulations exists, and their composition is selected depending on the biological material, its molecular cargo, and the anticipated route of application. For some EV-based preparations, a single protective compound may suffice, whereas other formulations require mixtures that perform several roles simultaneously. These mixtures may include buffering agents, osmotic regulators, fillers, or additional stabilizers [5, 18, 28]. Proper buffering is especially important: the pH of the medium must remain stable during freezing and drying. Certain common buffers, such as phosphate-based ones, experience significant and/or non-linear pH shifts in response to temperature changes and therefore can compromise vesicle integrity. Replacing them with temperature-stable buffers helps avoid undesirable osmotic shrinkage or swelling during the critical phases of freezing [17, 19-21]. Besides sugars, biopolymers such as alginate can serve as auxiliary stabilizers. For example, alginate enhances structural durability by interacting with protein surfaces and lipid membranes. It also reduces formation of nucleation sites and supports the stability of exosomes during the freeze-drying cycle.

Among all lyoprotective agents, disaccharides are considered the most versatile. They substitute the hydration shell around vesicles by forming extensive hydrogen-bond networks with phospholipid headgroups. As a result, amorphous sugar matrix is produced which prevents aggregation, fusion, or denaturation of proteins within EVs [8, 11, 22, 45]. Trehalose is widely recognized as the most effective of these disaccharides for preserving extracellular vesicles and other delicate biological systems during lyophilization [6, 13, 17, 21]. Found naturally in yeast, fungi, bacteria, and many plants, trehalose enables these organisms to withstand severe dehydration or freezing. During lyophilization of EVs, trehalose acts as a molecular scaffold that maintains vesicle conformation as the water content decreases. Its efficacy has been demonstrated across many biological preparations, including cell-free products and multiple cell types [22, 24, 40].

Trehalose is frequently combined with membrane-permeant cryoprotectants such as 10% DMSO. This dual approach often improves post-thaw recovery, enhances proliferative capacity of thawed cells, and preserves membrane integrity and intercellular contacts when multicellular constructs are being cryopreserved. Experimental findings point to an optimal working range of 100-400 mM trehalose. Higher concentrations of trehalose can compromise viability through osmotic stress. Interestingly, introducing trehalose enables the use of lower DMSO concentrations, which is particularly beneficial considering the cytotoxic potential of high DMSO levels. The natural origin, affordability, and broad compatibility of trehalose make it a preferred choice for both cellular and acellular products.

Evidence accumulated over recent years clearly shows that trehalose-based formulations maintain the structure and biological activity of EVs over long periods of storage [18, 24, 25, 34]. Because lyophilization inevitably exposes samples to physical, mechanical, and chemical stresses, the strategic use of lyoprotectants – whether individually or in combination – is essential for minimizing molecular degradation, preserving vesicle morphology, and maintaining functional characteristics during all stages of the freeze-dry process [11, 13, 46, 47].

5. Discussion

Contemporary experimental evidence highlights the broad potential of exosomes and EVs in regenerative medicine, dermatological applications, oncological therapy, and other biomedical areas. It becomes a valuable tool for anti-tumor and regenerative therapy among other well-documented approaches including application of engineered tissues and genetic modification of target cells [49]. Their functional capacities and therapeutic roles are summarized in Figure 2 (see the next page).

In the context of EVs production and preservation, the advantages of lyophilization include minimal chemical deterioration during storage, the formation of a dry, stable formulation, improved sterility management, and the option for transport and storage under non-refrigerated conditions. On the other hand, successful freeze-drying requires continuous monitoring of several critical factors. Foremost among them the shelf temperature during processing, the pressure profile within the sublimation chamber, and the duration and sequence of each phase of the freeze-dry cycle. Additional factors and actions, such as maintaining an optimal pH range, incorporating thermal stabilizers, and selecting appropriate formulation additives, are also necessary to maintain the functional integrity of sensitive EV components. It is particularly important to ensure that the final lyophilized product contains a controlled level of residual moisture, as this parameter directly affects long-term stability, physical robustness, and resistance to degradation, which is crucial for membrane-bound vesicles. A properly designed freeze-drying procedure allows EVs to remain in a dry, solid state without losing their biological properties, even when kept at ambient temperature. Studies demonstrate that exosomes preserved with trehalose maintain key activities, including anti-inflammatory and anti-fibrotic effects, at levels comparable to freshly isolated preparations [48]. Nonetheless, several challenges remain unresolved. For example, there is still no universally accepted standard for evaluating the quality of lyophilized EVs, nor is there consensus on how freeze-drying

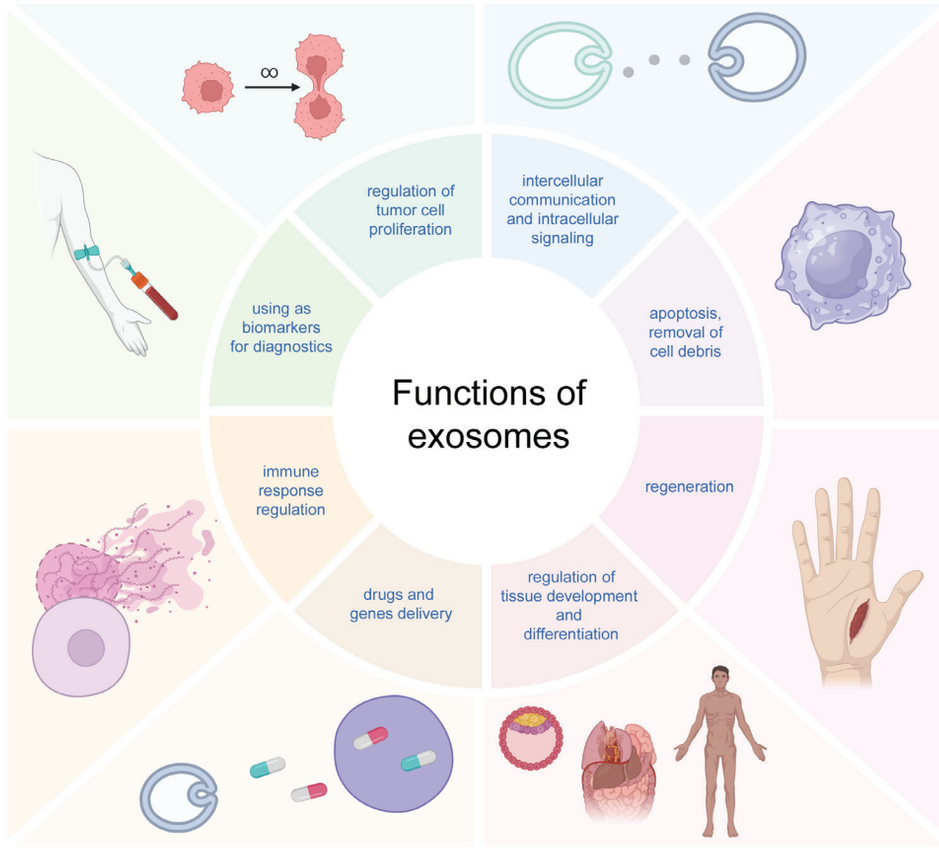


Figure 2 – Key functional features of exosomes in cell therapy. Image was created in Biorender.com

Image was created in Biorender.com.

may influence the specificity of their biological action or alter mechanisms underlying their therapeutic effects.

Current research aims to address these issues and improve the suitability of freeze-dried products for pharmaceutical and clinical applications, an effort that remains crucial for the advancement of EV-based therapies [50]. In this context, one of the most promising approaches is the implementation of AI-based algorithms and tools for various aspects related to the production, storage, and infusion of therapeutic EVs, similar to the AI-based approaches that are already in action in other fields of medicine [51]. Currently, AI is considered a powerful tool to design EVs with enhanced specificity to the target cells, e.g. due to the attaining highly precise delivery to the tissues [52]. Specifically, AI-based tools are used to characterize exosomes for their utilization in anti-cancer treatments as well as in other diseases [53]. It may also include the development of AI-assisted models for simulation of thermal changes in structural properties of the EVs during cryopreservation and dry-freeze, physicochemical stability of their cargo composition, and how the membrane-fusion ability of exosomes can be affected by the environmental changes like pH, temperature, or concentration of medium components. These AI-assisted approaches may facilitate designing and optimizing robust protocols of production and storage of EVs taking into account the type of original cells, culture medium composition, and the needed quantity of final cell-free product.

Conclusion

Lyophilization is increasingly recognized as a promising and technically mature approach for stabilizing cell-free biological therapeutics, particularly exosomes and other extracellular vesicles. By this approach, it is possible to create product that can be stored and transported for extended periods without reliance on deep-freeze conditions and, importantly, without losing its functional quality. When the process is supplemented with carefully selected cryo- and lyophiloprotectants and finely tuned operational parameters, the essential structural attributes and biological functions of these nanoscale vesicles – including membrane architecture, molecular cargo, and bioactivity – can be reliably preserved. At the same time, specific limitations persist, such as the length of the lyophilization cycle, the necessity of sterile reconstitution before use, and the dependence on lyoprotectants to maintain EVs activity at levels comparable to native vesicles. Furthermore, the regulatory classification and approval of lyophilized EV-based products as medicinal agents remain in early stages and require clearer legislative frameworks. For these reasons, progress in this field depends on coordinated interdisciplinary efforts. Unified actions by basic and applied researchers, engineering solutions for reproducible manufacturing, and implementation of validated quality-assurance systems constitute the potential for advancements. Such developments are essential for establishing a new class of stable, clinically applicable cell-free biopharmaceuticals.

Author Contributions: A.B., E.C., A.Zh., A.A., and A.G. – conception of the work; A.Zh., A.A. – reviewing literature, writing the draft; A.Zh. – visualization; A.B., E.C., A.A., A.Zh., and A.G. – revision of final version. All authors reviewed and approved the final manuscript for submission.

Disclosures: There is no conflict of interest for all authors.

Acknowledgements: None.

Funding: This work was supported by the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. BR28713159).

Data availability statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

AI Usage Statement: Not applicable.

References

1. Kumar MA, Baba SK, Sadida HQ, Marzooqi SA, Jerobin J, Altemani FH, Algehainy N, Alanazi MA, Abou-Samra AB, Kumar R, Al-Shabeeb Akil AS, Macha MA, Mir R, Bhat AA. Extracellular vesicles as tools and targets in therapy for diseases. *Signal Transduct Target Ther.* 2024;9(1):27. https://doi.org/10.1038/s41392-024-01735-1
2. Cheng L, Hill AF. Therapeutically harnessing extracellular vesicles. *Nat Rev Drug Discov.* 2022;21:379-399. https://doi.org/10.1038/s41573-022-00410-w
3. Bahr MM, Amer MS, Abo-El-Sooud K, Abdallah AN, El-Tookhy OS. Preservation techniques of stem cells extracellular vesicles: a gate for manufacturing of clinical grade therapeutic extracellular vesicles and long-term clinical trials. *Int J Vet Sci Med.* 2020;8(1):1-8. https://doi.org/10.1080/23144599.2019.1704992
4. Susa F, Limongi T, Borgione F, Peiretti S, Vallino M, Cauda V, Pisano R. Comparative Studies of Different Preservation Methods and Relative Freeze-Drying Formulations for Extracellular Vesicle Pharmaceutical Applications. *ACS Biomater Sci Eng.* 2023;9(10):5871-5885. https://doi.org/10.1021/acsbiomaterials.3c00678
5. Yuan F, Li YM, Wang Z. Preserving extracellular vesicles for biomedical applications: consideration of storage stability before and after isolation. *Drug Deliv.* 2021;28(1):1501-1509. https://doi.org/10.1080/10717544.2021.1951896
6. Romanov YA, Volgina NE, Dugina TN, Kabaeva NV, Sukhikh GT. Effect of Storage Conditions on the Integrity of Human Umbilical Cord Mesenchymal Stromal Cell-Derived Microvesicles. *Bull Exp Biol Med.* 2019;167(1):131-135. https://doi.org/10.1007/s10517-019-04476-2
7. Shao S, Fang H, Li Q, Wang G. Extracellular vesicles in Inflammatory Skin Disorders: from Pathophysiology to Treatment. *Theranostics*. 2020;10(22):9937-9955. https://doi.org/10.7150/thno.45488
8. Latipbayeva M, Askarov M, Seisembekov T, Baigenzhin A, Kupenov B, Tolegenuly A, Zakarina D, Mansharipova A, Akhayeva A, Kussaiynkhan G. Study of the Efficacy of Autologous Cell Therapy in Dilated Cardiomyopathy. *J Clin Med Kaz*. 2025;22(4):37-46. https://doi.org/10.23950/jcmk/16589
9. Yeltokova M, Akhmedyanova Z, Bayanbayeva Z, Khassenova A, Yermekova K. The use of autologous mesenchymal stem cells in complications of diabetes mellitus, in particular diabetic retinopathy: inputs and insights. *J Clin Med Kaz.* 2022;19(2):9-13. https://doi.org/10.23950/jcmk/11938
10. Nooshabadi VT, Mardpour S, Yousefi-Ahmadipour A, Allahverdi A, Izadpanah M, Daneshimehr F, Ai J, Banafshe HR, Ebrahimi-Barough S. The extracellular vesicles-derived from mesenchymal stromal cells: A new therapeutic option in regenerative medicine. *J Cell Biochem.* 2018;119(10):8048-8073. https://doi.org/10.1002/jcb.26726
11. Ahmadian S, Jafari N, Tamadon A, Ghaffarzadeh A, Rahbarghazi R, Mahdipour M. Different storage and freezing protocols for extracellular vesicles: a systematic review. *Stem Cell Res Ther.* 2024;15(1):453. https://doi.org/10.1186/s13287-024-04005-7
12. El Baradie KBY, Nouh M, O'Brien Iii F, Liu Y, Fulzele S, Eroglu A, Hamrick MW. Freeze-Dried Extracellular Vesicles From Adipose-Derived Stem Cells Prevent Hypoxia-Induced Muscle Cell Injury. *Front Cell Dev Biol.* 2020;8:181. https://doi.org/10.3389/fcell.2020.00181
13. Trenkenschuh E, Richter M, Heinrich E, Koch M, Fuhrmann G, Friess W. Enhancing the Stabilization Potential of Lyophilization for Extracellular Vesicles. *Adv Healthc Mater.* 2022;11(5):e2100538. https://doi.org/10.1002/adhm.202100538
14. Lei X, Ring S, Jin S, Singh S, Mahnke K. Extracellular Vesicles and Their Role in Skin Inflammatory Diseases: From Pathogenesis to Therapy. *Int J Mol Sci.* 2025;26(8):3827. https://doi.org/10.3390/ijms26083827
15. Li Z, Yan J, Li X, Chen H, Lin C, Zhang Y, Gao T, Zhang Y, Shu Y, Pan S, Zhang Y. Advancements in extracellular vesicles biomanufacturing: a comprehensive overview of large-scale production and clinical research. *Front Bioeng Biotechnol.* 2025;13:1487627. https://doi.org/10.3389/fbioe.2025.1487627
16. Yu H, Feng H, Zeng H, Wu Y, Zhang Q, Yu J, Hou K, Wu M. Exosomes: The emerging mechanisms and potential clinical applications in dermatology. *Int J Biol Sci.* 2024;20(5):1778-1795. https://doi.org/10.7150/ijbs.92897
17. Cardoso RMS, Rodrigues SC, Gomes CF, Duarte FV, Romao M, Leal EC, Freire PC, Neves R, Simões-Correia J. Development of an optimized and scalable method for isolation of umbilical cord blood-derived small extracellular vesicles for future clinical use. *Stem Cells Transl Med.* 2021;10(6):910-921. https://doi.org/10.1002/sctm.20-0376
18. Izutsu KI. Applications of Freezing and Freeze-Drying in Pharmaceutical Formulations. *Adv Exp Med Biol.* 2018;1081:371-383. https://doi.org/10.1007/978-981-13-1244-1_20
19. Guarro M, Suñer F, Lecina M, Borrós S, Fornaguera C. Efficient extracellular vesicles freeze-dry method for direct formulations preparation and use. *Colloids Surf B Biointerfaces*. 2022;218:112745. https://doi.org/10.1016/j.colsurfb.2022.112745
20. Centner CS, Murphy EM, Priddy MC, Moore JT, Janis BR, Menze MA, DeFilippis AP, Kopechek JA. Ultrasound-induced molecular delivery to erythrocytes using a microfluidic system. *Biomicrofluidics*. 2020;14(2):024114. https://doi.org/10.1063/1.5144617
21. Liu Q, Wu J, Wang H, Jia Z, Li G. Human Infrapatellar Fat Pad Mesenchymal Stem Cell-derived Extracellular Vesicles Purified by Anion Exchange Chromatography Suppress Osteoarthritis Progression in a Mouse Model. *Clin Orthop Relat Res.* 2024;482(7):1246-1262. https://doi.org/10.1097/CORR.0000000000003067

22. Kusuma GD, Barabadi M, Tan JL, Morton DAV, Frith JE, Lim R. To Protect and to Preserve: Novel Preservation Strategies for Extracellular Vesicles. *Front Pharmacol.* 2018;9:1199. https://doi.org/10.3389/fphar.2018.01199
23. Golan ME, Stice SL. Extracellular vesicle lyophilization for enhanced distribution to the point of care. *Extracellular Vesicle,* 2024;3:100041. https://doi.org/10.1016/j.vesic.2024.100041
24. Kang WY, Shin EK, Kim EH, Kang MH, Bang CY, Bang OY, Cha JM. Lyoprotectant Constituents Suited for Lyophilization and Reconstitution of Stem-Cell-Derived Extracellular Vesicles. *Biomater Res.* 2024;28:0005. https://doi.org/10.34133/bmr.0005
25. Manchon E, Hirt N, Bouaziz JD, Jabrane-Ferrat N, Al-Daccak R. Stem Cells-Derived Extracellular Vesicles: Potential Therapeutics for Wound Healing in Chronic Inflammatory Skin Diseases. *Int J Mol Sci.* 2021;22(6):3130. https://doi.org/10.3390/ijms22063130
26. Bari E, Perteghella S, Di Silvestre D, Sorlini M, Catenacci L, Sorrenti M, Marrubini G, Rossi R, Tripodo G, Mauri P, Marazzi M, Torre ML. Pilot Production of Mesenchymal Stem/Stromal Freeze-Dried Secretome for Cell-Free Regenerative Nanomedicine: A Validated GMP-Compliant Process. *Cells.* 2018;7(11):190. https://doi.org/10.3390/cells7110190
27. Rodrigues SC, Cardoso RMS, Freire PC, Gomes CF, Duarte FV, Neves RPD, Simões-Correia J. Immunomodulatory Properties of Umbilical Cord Blood-Derived Small Extracellular Vesicles and Their Therapeutic Potential for Inflammatory Skin Disorders. *Int J Mol Sci.* 2021;22(18):9797. https://doi.org/10.3390/ijms22189797
28. Miyamoto Y, Ikeuchi M, Noguchi H, Hayashi S. Long-term Cryopreservation of Human and other Mammalian Cells at -80 °C for 8 Years. *Cell Med.* 2018;10:2155179017733148. https://doi.org/10.1177/2155179017733148
29. Bajerski F, Nagel M, Overmann J. Microbial occurrence in liquid nitrogen storage tanks: a challenge for cryobanking? *Appl Microbiol Biotechnol.* 2021;105(20):7635-7650. https://doi.org/10.1007/s00253-021-11531-4
30. Lyu N, Knight R, Robertson SYT, Dos Santos A, Zhang C, Ma C, Xu J, Zheng J, Deng SX. Stability and Function of Extracellular Vesicles Derived from Immortalized Human Corneal Stromal Stem Cells: A Proof of Concept Study. *AAPS J*. 2022;25(1):8. https://doi.org/10.1208/s12248-022-00767-1
31. van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of extracellular vesicles. *Nat Rev Mol Cell Biol*. 2018;19(4):213-228. https://doi.org/10.1038/nrm.2017.125
32. Jeppesen DK, Fenix AM, Franklin JL, Higginbotham JN, Zhang Q, Zimmerman LJ, Liebler DC, Ping J, Liu Q, Evans R, Fissell WH, Patton JG, Rome LH, Burnette DT, Coffey RJ. Reassessment of Exosome Composition. *Cell*. 2019;177(2):428-445.e18. https://doi.org/10.1016/j.cell.2019.02.029
33. Pluchino S, Smith JA. Explicating Exosomes: Reclassifying the Rising Stars of Intercellular Communication. *Cell.* 2019;177(2):225-227. https://doi.org/10.1016/j.cell.2019.03.020
34. Meldolesi J. Exosomes and Ectosomes in Intercellular Communication. *Curr Biol.* 2018;28(8):R435-R444. https://doi.org/10.1016/j.cub.2018.01.059
35. You B, Yang Y, Zhou Z, Yan Y, Zhang L, Jin J, Qian H. Extracellular Vesicles: A New Frontier for Cardiac Repair. *Pharmaceutics*. 2022;14(9):1848. https://doi.org/10.3390/pharmaceutics14091848
36. Omoto ACM, do Carmo JM, da Silva AA, Hall JE, Mouton AJ. Immunometabolism, extracellular vesicles and cardiac injury. *Front Endocrinol (Lausanne).* 2024;14:1331284. https://doi.org/10.3389/fendo.2023.1331284
37. Monroe DG, Javeed N, Rowsey JL, Ruan M, McCabe CE, Piatkowski BT, Roy A, Bobbili MR, Grillari J, Khosla S. Isolation and characterization of bone mesenchymal cell small extracellular vesicles using a novel mouse model. *J Bone Miner Res.* 2024;39(11):1633-1643. https://doi.org/10.1093/jbmr/zjae135
38. Urabe F, Kosaka N, Ito K, Kimura T, Egawa S, Ochiya T. Extracellular vesicles as biomarkers and therapeutic targets for cancer. *Am J Physiol Cell Physiol*. 2020;318(1):C29-C39. https://doi.org/10.1152/ajpcell.00280.2019
39. Buratta S, Tancini B, Sagini K, Delo F, Chiaradia E, Urbanelli L, Emiliani C. Lysosomal Exocytosis, Exosome Release and Secretory Autophagy: The Autophagic- and Endo-Lysosomal Systems Go Extracellular. *Int J Mol Sci.* 2020;21(7):2576. https://doi.org/10.3390/ijms21072576
40. Budgude P, Kale V, Vaidya A. Cryopreservation of mesenchymal stromal cell-derived extracellular vesicles using trehalose maintains their ability to expand hematopoietic stem cells in vitro. Cryobiology. 2021;98:152-163. https://doi.org/10.1016/j.cryobiol.2020.11.009
41. Amari L, Germain M. Mitochondrial Extracellular Vesicles - Origins and Roles. *Front Mol Neurosci.* 2021;14:767219. https://doi.org/10.3389/fnmol.2021.767219
42. Liu Q, Li S, Dupuy A, Mai HL, Sailliet N, Logé C, Robert JH, Brouard S. Exosomes as New Biomarkers and Drug Delivery Tools for the Prevention and Treatment of Various Diseases: Current Perspectives. *Int J Mol Sci.* 2021;22(15):7763. https://doi.org/10.3390/ijms22157763
43. Hosseini M, Roshangar L, Raeisi S, Ghahremanzadeh K, Negargar S, Tarmahi V, Hosseini V, Raeisi M, Rahimi E, Ebadi Z. The Therapeutic Applications of Exosomes in Different Types of Diseases: A Review. *Curr Mol Med*. 2021;21(2):87-95. https://doi.org/10.2174/1566524020666200610164743
44. Zhang F, Lu X, Zhu X, Yu Z, Xia W, Wei X. Real-time monitoring of small extracellular vesicles (sEVs) by in vivo flow cytometry. *J Extracell Vesicles.* 2024;13(10):e70003. https://doi.org/10.1002/jev2.70003
45. Wang Y, Yuan S, Zhou L, Yang K, Jin Z, Lin A, Yang C, Tian W. Cutting-Edge Progress in the Acquisition, Modification and Therapeutic Applications of Exosomes for Drug Delivery. *Int J Nanomedicine.* 2025;20:5059-5080. https://doi.org/10.2147/IJN.S516840
46. Qazi REM, Sajid Z, Zhao C, Hussain I, Iftikhar F, Jameel M, Rehman FU, Mian AA. Lyophilization Based Isolation of Exosomes. *Int J Mol Sci.* 2023;24(13):10477. https://doi.org/10.3390/ijms241310477
47. Assegehegn G, Brito-de la Fuente E, Franco JM, Gallegos C. The Importance of Understanding the Freezing Step and Its Impact on Freeze-Drying Process Performance. *J Pharm Sci.* 2019;108(4):1378-1395. https://doi.org/10.1016/j.xphs.2018.11.039
48. Jafari A, Mirzaei Y, Mer AH, Rezaei-Tavirani M, Jafari Z, Niknejad H. Comparison of the effects of preservation methods on structural, biological, and mechanical properties of the human amniotic membrane for medical applications. *Cell Tissue Bank.* 2024;25(1):305-323. https://doi.org/10.1007/s10561-023-10114-z
49. Aijaz M, Ahmad S, Alam S. Regenerative Medicine Unveiled: Principles, Technologies, and Clinical Breakthroughs in Tissue Regeneration. *J Clin Med Kaz.* 2025;22(5):105-118. https://doi.org/10.23950/jcmk/16880

50. Alexeenko AA, Darwish A, Strongrich D, Kazarin P, Patil C, Tower CW, Wheeler IS, Munson E, Zhou Q, Narsimhan V, Yoon K, Nail SL, Cofer A, Stanbro J, Renawala H, Roth D, DeMarco F, Griffiths J, Peroulis D. Randomized-field microwave-assisted pharmaceutical lyophilization with closed-loop control. *Sci Rep.* 2025;15(1):10536. https://doi.org/10.1038/s41598-025-91642-4
51. Oriaku I, Okechukwu O, Okeoma OI, Gab-Obinna C, Bala JI, Adejumobi AM, Awoyomi OO. Artificial Intelligence in Transfusion Medicine: Current Applications, Opportunities, and Challenges. *Epidemiology and Health Data Insights.* 2025;1(6):ehdi023. https://doi.org/10.63946/ehdi/17470
52. Greenberg ZF, Graim KS, He M. Towards artificial intelligence-enabled extracellular vesicle precision drug delivery. *Adv Drug Deliv Rev.* 2023;199:114974. https://doi.org/10.1016/j.addr.2023.114974
53. Picchio V, Pontecorvi V, Dhori X, Bordin A, Floris E, Cozzolino C, Frati G, Pagano F, Chimenti I, De Falco E. The emerging role of artificial intelligence applied to exosome analysis: from cancer biology to other biomedical fields. *Life Sci.* 2025;375:123752. https://doi.org/10.1016/j.lfs.2025.123752

CONTENTS

(E-ISSN 2313-1519)
JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN

Mukhit Dossov, Serik Seitenov, Nazira Bekenova, Galymzhan Kuatbay
THE INFLUENCE OF DISEASE DURATION IN PHARMACORESISTANT FOCAL EPILEPSY AND TYPE OF ANTIEPILEPTIC THERAPY ON INTRAOPERATIVE INFUSION THERAPY PARAMETERS **4**

Maja Avramovska, Petar Avramovski, Liljana Todorovska, Biljana Taleva, Kosta Sotiroski, Dejan Zdraveski
THE SILENT THREATS OF AGING: FEAR OF FRAGILE BONES, STIFF ARTERIES, OR TIME'S INEVITABLE BETRAYAL? AN AGE-MATCHED STUDY **10**

Meltem Saygili, Aysun Kandemir Ture, Nesrin Akca
HEALTH CARE QUALITY AND PATIENT SATISFACTION IN HEMODIALYSIS CLINIC: INVESTIGATION OF PATIENT-REPORTED OUTCOMES **19**

Cansu Yaşar, Şeyma Zehra Altunkürek
THE EFFECT OF TRANSTHEORETICAL MODEL-BASED ONLINE HEALTH EDUCATION ON NUTRITION AND PHYSICAL ACTIVITY CHANGE PROCESSES IN OVERWEIGHT UNIVERSITY STUDENTS: A RANDOMIZED CONTROLLED TRIAL **28**

Emmanuel Jimmy, Obeten Ekpo, Bassey Ekeng, Chidimma Ahaneku, Bernard Bassey, Simon Ereh, Thelma Bassey, L ove Okafor, Agbo James, Eshan Henshaw, Chibueze Njoku, Etete Petters
EVALUATION OF RESPIRATORY SYMPTOMS AND VENTILATORY FUNCTION AMONGST FOOD VENDORS IN CALABAR, NIGERIA: A CROSS-SECTIONAL COMPARATIVE STUDY **37**

Alapati Lakshmi Rushitha, Dinesh Kamaraj, Rangasamy Krishnamoorthi, Panneerselvam Periasamy, Arbind Kumar Choudhary
DIAGNOSTIC PERFORMANCE OF THE STEPSS SCORE AND MACHINE LEARNING MODELS FOR PREDICTING OUTCOMES IN PEDIATRIC STATUS EPILEPTICUS: A PROSPECTIVE OBSERVATIONAL STUDY **44**

Ashini A. Patel, Caitlyn R. Curley, Ganesh Chilukuri, Hannah Lee, Samantha Snyder, Vivek Joshi
POLYCYSTIC OVARY SYNDROME IN MEDICAL STUDENTS: A CROSS-SECTIONAL STUDY OF THE ACADEMIC AND PSYCHOSOCIAL IMPACT **54**

Aya S. Aldewachi, Mohammed I. Aladul
THE IMPACT OF GLUCAGON-LIKE PEPTIDE-1 RECEPTOR AGONISTS ON OBESE INDIVIDUALS' QUALITY OF LIFE AND EATING BEHAVIORS **60**

Hariharan Venkatasubramanian, Vijayasamundeeswari Chinnathambipalayam Kandaswamy
RAPID HEPATIC FUNCTION IMPROVEMENT WITH CULTURALLY ADAPTED LOW-CARBOHYDRATE PALEOLITHIC DIET IN SOUTH ASIAN ADULTS WITH METABOLIC DYSFUNCTION: A 12-WEEK PROSPECTIVE STUDY WITH PREDICTIVE MODELING **73**

Bakhyt M. Sultanbekova, Marzhan Brimzhanova, Gulnara Kandygulova, Nurgul Alekenova,Anuar Akhmetzhan, Natalya Pirozhenko
RISK MANAGEMENT IN GYNECOLOGICAL DEPARTMENTS OF HOSPITALS IN ALMATY **80**

Beibit Abdikenov, Nurbek Saidnassim, Birzhan Ayanbayev, Aruzhan Imasheva
CLASSIFICATION OF COMBINED MLO AND CC MAMMOGRAPHIC VIEWS USING VISION–LANGUAGE MODELS **89**

Ruslan Akhmedullin, Zhandos Burkitbayev, Tair Koishibayev, Zhanat Spatayev, Abylaikhan Sharmenov, Oxana Shatkovskaya, Dinara Zharlyganova, Almira Manatova,Zhuldyz Kuanysh, Sanzhar Shalekenov
SURGERY AND SURVIVAL IN PATIENTS WITH INTRAHEPATIC CHOLANGIOCARCINOMA: A BAYESIAN INFERENCE **96**

Altynshash Jaxybayeva, Dina Kalinina
UNDERSTANDING THE HEALTHCARE LANDSCAPE FOR PEDIATRIC EPILEPSY PATIENTS IN KAZAKHSTAN: CHALLENGES AND OPPORTUNITIES **103**

Olabisi Promise Lawal, Aliyu Olanrewaju Olaniyi, Uche Cyprain Tochukwu, Catherine Rono, Mayowa Josiah Badmus, Misbahudeen Olohuntoyin Abdulkareem, Udeichi Mmesoma Odinaka[7], Idris Olumide Orenolu, Udeichi Nnenna Chidimma
MUCOSAL IMMUNE MARKERS AND THEIR IMPACT ON ANTIMICROBIAL RESISTANCE: A SYSTEMATIC REVIEW **109**

Abay Baigenzhin, Elmira Chuvakova, Aigerim Zhakupova, Aizhan Akhayeva, Anastassiya Ganina
EXTRACELLULAR VESICLES AND LYOPHILIZATION: GETTING FUNCTIONALLY STABLE CELL-FREE BIOMEDICAL PRODUCTS **117**