Acıbadem Univ. Sağlık Bilim. Derg. 17 (October, November, December) 2026
https://doi.org/10.31067/acusaglik.2009917

ORIGINAL ARTICLE / ARAŞTIRMA YAZISI

Psychiatry/Psikiyatri

Renal and Hepatic Biochemical Trajectories During Lithium and Valproate Treatment in Bipolar Disorder

İlker Güneysu[1], Sare Aydın[1], Burcu Eser[1]

[1] Department of Psychiatry, Faculty of Medicine, Tokat Gaziosmanpaşa University, Tokat, Türkiye

ABSTRACT

Purpose: To compare within-person renal and hepatic biochemical trajectories during lithium and valproate treatment in bipolar disorder and determine whether longitudinal slopes differed between treatment groups.

Methods: This single-center retrospective cohort included 96 patients with bipolar disorder treated with lithium (n=35) or valproate (n=61), contributing 282 laboratory records. Random-intercept linear mixed-effects models for creatinine, blood urea nitrogen (BUN), aspartate aminotransferase (AST), and alanine aminotransferase (ALT) included age at the first available laboratory record, sex, time, treatment group, and the time × treatment interaction. Outcome-specific adjusted models comprised 275-278 observations from 95 patients. For creatinine, a treatment-specific piecewise model with a 6-month knot was supplemented by data-quality and random-slope sensitivity analyses.

Results: Mean follow-up was 312.22 ± 199.21 days. The time × treatment interaction was not statistically significant for creatinine (B=-0.0019, p=.303), BUN (B=-0.0853, p=.160), AST (B=-0.0070, p=.285), or ALT (B=-0.0059, p=.442). Male sex was associated with higher creatinine, AST, and ALT levels, whereas older baseline age was associated with higher BUN. In the piecewise model, neither early nor late slopes differed by treatment (joint Wald $\chi^2(2)=0.544$, p=.762). All sensitivity analyses supported the primary finding.

Conclusion: No statistically detectable difference was observed between lithium and valproate groups in the longitudinal slopes of the evaluated markers. The findings do not establish equivalence or long-term safety; average short- to medium-term trajectories did not diverge under regular, individualized laboratory monitoring.

Keywords: Bipolar disorder; lithium; valproic acid (valproate); kidney function tests; liver function tests

ÖZET

Amaç: Bipolar bozuklukta lityum ve valproat tedavisi sırasında renal ve hepatik biyokimyasal göstergelerin kişi içi seyrini ve tedavi grupları arasındaki zamansal eğim farklarını karşılaştırmaktır.

Yöntem: Tek merkezli retrospektif kohortta bipolar bozukluk tanılı, lityum (n=35) veya valproat (n=61) kullanan 96 hastaya ait 282 laboratuvar kaydı değerlendirildi. Kreatinin, kan üre azotu (BUN), aspartat aminotransferaz (AST) ve alanin aminotransferaz (ALT) için ilk laboratuvar kaydındaki yaş, cinsiyet, zaman, tedavi grubu ve zaman × tedavi etkileşimini içeren rastgele kesişimli doğrusal karma etkili modeller kuruldu. Sonuca özgü ayarlı modeller 95 hastadan 275-278 gözleme dayanıyordu. Kreatinin için tedaviye özgü erken ve geç eğimleri içeren altı ay düğüm noktalı parçalı model ile veri kalitesi ve rastgele eğim duyarlılık analizleri yapıldı.

Bulgular: Ortalama izlem süresi 312.22 ± 199.21 gündü. Zaman × tedavi etkileşimi kreatinin (B=-0.0019, p=.303), BUN (B=-0.0853, p=.160), AST (B=-0.0070, p=.285) ve ALT (B=-0.0059, p=.442) için anlamlı değildi. Erkek cinsiyet daha yüksek kreatinin, AST ve ALT düzeyleriyle; başlangıç yaşı daha yüksek BUN düzeyiyle ilişkiliydi. Parçalı modelde erken veya geç eğimler tedavi gruplarına göre farklılaşmadı (ortak Wald $\chi^2(2)=0.544$, p=.762). Tüm duyarlılık analizleri temel sonucu korudu.

Sonuç: Lityum ve valproat gruplarında değerlendirilen renal ve hepatik göstergelerin gözlenen izlem boyunca eğimleri arasında istatistiksel olarak saptanabilir bir fark bulunmadı. Bulgular eşdeğerlik veya uzun dönem güvenlilik kanıtı olarak değil, düzenli ve bireyselleştirilmiş laboratuvar izlemi altında kısa ve orta dönem ortalama seyirlerin ayrışmadığını gösteren gerçek yaşam verisi olarak yorumlanmalıdır.

Anahtar Kelimeler: Bipolar bozukluk; lityum; valproat; böbrek fonksiyon testleri; karaciğer fonksiyon testleri

İlker GÜNEYSU
0000-0003-4777-4524
Sare AYDIN
0000-0002-5081-5983
Burcu ESER
0000-0001-8609-3161

Correspondence: İlker Güneysu
Department of Psychiatry, Faculty of Medicine, Tokat Gaziosmanpaşa University, Tokat, Türkiye
Phone: +90 537 345 10 38
E-mail: ilkerguneysu@gmail.com

Received: 03.08.2026
Accepted: 07.09.2026

Copyright © 2026 the Author(s). Published by Acibadem University. This is an open access article licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND 4.0) International License, which is downloadable, re-usable and distributable in any medium or format in unadapted form and for noncommercial purposes only where credit is given to the creator and publishing journal is cited properly. The work cannot be used commercially without permission from the journal.

Bipolar disorder is a chronic psychiatric condition characterized by recurrent depressive and manic or hypomanic episodes and is associated with impaired functioning and an increased risk of premature mortality (1). The principal goals of treatment are to control acute mood episodes, prevent recurrence, and preserve long-term functioning. In this context, lithium and valproate are among the most widely used mood stabilizers for the acute and maintenance treatment of bipolar disorder (2,3).

Although the efficacy of lithium in reducing recurrence and suicide risk is well established, its narrow therapeutic range and potential effects, particularly on renal and thyroid function, require regular biochemical monitoring (4-6). Valproate is an important treatment option, especially for presentations with mixed features or rapid cycling, but it may be associated with aminotransferase elevations, metabolic changes, and, less commonly, severe hepatic toxicity (7). Safety assessment is therefore integral to clinical decision-making from treatment selection through maintenance therapy.

Registry and cohort studies directly comparing lithium and valproate have not produced entirely consistent findings for renal events. Some studies have reported stronger associations between lithium and chronic kidney disease or other renal outcomes, whereas more recent real-world comparisons have found smaller risk differences (8-11). Much of this literature has emphasized event-based outcomes such as chronic kidney disease, acute kidney injury, or treatment discontinuation (12,13), while routine clinical monitoring studies have often focused on whether recommended laboratory assessments are performed rather than directly comparing treatment-specific within-person biochemical slopes (14).

Mixed-effects models for repeated measurements can accommodate patients with different numbers and intervals of observations and can estimate baseline level differences and longitudinal change within the same statistical framework. In this observational setting, however, such modeling does not eliminate baseline confounding or permit causal attribution of longitudinal differences to treatment. This study aimed to compare the within-person trajectories of creatinine, blood urea nitrogen (BUN), aspartate aminotransferase (AST), and alanine aminotransferase (ALT) in patients with bipolar disorder treated with lithium or valproate and to examine whether longitudinal slopes differed between treatment groups. For creatinine, slopes during the first 6 months of observation and the subsequent period were also evaluated in an exploratory analysis.

Material and Methods

Study Design and Participants

This single-center retrospective cohort study was conducted at the Department of Psychiatry, Faculty of Medicine, Tokat Gaziosmanpaşa University, Tokat, Türkiye, between 2023 and 2025. Electronic medical records were reviewed for patients diagnosed with bipolar disorder according to DSM-5 criteria by a psychiatrist and monitored during treatment with lithium or valproate. The institutional screening included 664 patient records. Clinical eligibility screening was performed at the electronic-record review stage, before construction of the analytic laboratory dataset, by applying the predefined inclusion and exclusion criteria described below. After this screening, 96 patients were included in the analytic dataset (35 receiving lithium and 61 receiving valproate), meaning that 568 screened records did not enter the analytic cohort. The historical screening file did not retain criterion-specific exclusion counts; therefore, the 568 nonincluded records cannot be validly partitioned among individual exclusion reasons retrospectively.

Eligibility required at least three clinical follow-up visits, renal and hepatic laboratory assessments obtained as part of treatment monitoring, and use of either lithium or valproate, but not both. Patients with chronic renal failure, chronic liver disease, or concurrent lithium and valproate use were excluded. The ≥3-visit criterion referred to clinical follow-up, whereas the analytic record count referred to dated laboratory panels available in the extracted dataset. Laboratory information was not complete across all clinical visits or all analytes; consequently, after extraction, individual patients contributed one to four usable laboratory records. A usable record was defined as a dated laboratory panel contributing at least one of the four target analytes (creatinine, BUN, AST, or ALT), with outcome-specific missingness handled at the model level. Because the raw dataset included four patients with follow-up exceeding 18 months, these patients were retained in the primary analysis, and an analysis excluding them was conducted as a prespecified sensitivity assessment.

Data Source and Variables

Demographic characteristics, treatment group, and laboratory results were obtained from electronic medical records. Creatinine and BUN were evaluated as renal markers, and AST and ALT as hepatic markers. The raw analytic file contained 282 records from 96 patients. The first available laboratory measurement for each patient defined the start of observed laboratory follow-up, and the number of days until subsequent measurements was divided by 30 to create a time variable in months. Accordingly, the time axis represents time since the first available laboratory measurement, not time since lithium or valproate initiation.

Because age appeared to have been recalculated at each visit, baseline age from the patient's first record was used as a patient-level covariate instead of visit age, which would change mechanically over time. One patient with missing baseline age in all records was excluded from the adjusted models. Treatment was coded as lithium or valproate, and sex as female or male. Creatinine and BUN were expressed in mg/dL, and AST and ALT in U/L.

Statistical Analysis

Statistical analyses were performed using R version 4.5.1 within the RStudio environment. Continuous variables are presented as mean ± standard deviation or median (range), and categorical variables as number and percentage. Baseline continuous variables were compared using the Mann-Whitney U test, and sex using the Yates-corrected chi-square test. Hedges g for continuous variables and Cramér V for sex were calculated to quantify imbalance between the observational groups.

A linear mixed-effects model with a patient-specific random intercept was fitted for each outcome. Fixed effects included time in months, treatment group, baseline age centered around the overall mean, sex, and the time × treatment interaction. Lithium and female sex were the reference categories. Models were estimated using restricted maximum likelihood, and coefficients were reported with Wald standard errors, 95% confidence intervals, and two-sided Wald z tests. A log(x+1) transformation was applied to AST and ALT. The primary inference was based on the time × treatment interaction, which represents the difference in longitudinal slopes between treatment groups; nonsignificant interactions were not interpreted as evidence of equivalence.

Because of outcome-specific missingness, the models included 275-278 observations from 95 patients. The mixed-effects models were assumed to address missing outcome values under a missing-at-random assumption conditional on the observed covariates. Models with a random slope were fitted as sensitivity analyses for slope heterogeneity.

A secondary exploratory piecewise model with a knot at 6 months was fitted for creatinine. The 6-month cut-point was selected as a clinically interpretable early-monitoring window rather than estimated from the observed data. The model included early- and late-period slopes, treatment group, baseline age, sex, and the early period × treatment and late period × treatment interactions. Treatment-specific marginal slopes and a joint Wald test of the two interaction terms were reported. Seven laboratory-panel pairs were flagged as potentially duplicate or repeated measurements because, within the same patient, the target laboratory values were identical and the panels were recorded 7-8 days apart. Sensitivity analyses excluded the later record in each flagged pair, excluded four patients followed for more than 18 months, restricted the analysis to patients with at least three usable records, and applied all data-quality restrictions simultaneously. Statistical significance was defined as $p<.05$.

Ethics Approval

The study was approved by the Non-Interventional Scientific Research Ethics Committee of Tokat Gaziosmanpaşa Faculty of Medicine (Decision No: 25-MOBAEK-355; Date: 14.10.2025). Individual informed consent was not obtained because the study was based on a retrospective review of medical records.

Results

Sample and Data Structure

Patients in the lithium group were older than those in the valproate group (46.66 ± 11.20 vs 40.58 ± 13.83 years; p=.021; Hedges g=0.47), and the proportion of women was higher in the lithium group (71.4% vs 47.5%; p=.040; Cramér V=0.21). Baseline creatinine, BUN, AST, and ALT levels did not differ significantly between groups, and the standardized differences were small (Table 1).

Table 1. Baseline Demographic and Biochemical Characteristics by Treatment Group

Variable	Lithium	Valproate	p	Effect size
Baseline age, years	46.66 ± 11.20 (n=35)	40.58 ± 13.83 (n=60)	.021	Hedges g = 0.47
Female, n (%)	25 (71.4)	29 (47.5)	.040	Cramér V = 0.21
Creatinine, mg/dL	0.84 ± 0.19 (n=35)	0.80 ± 0.16 (n=59)	.247	Hedges g = 0.25
BUN, mg/dL	11.88 ± 3.83 (n=35)	12.19 ± 4.13 (n=60)	.766	Hedges g = -0.08
AST, U/L	20.65 ± 13.32 (n=34)	20.67 ± 11.34 (n=61)	.507	Hedges g = 0.00
ALT, U/L	20.55 ± 14.33 (n=32)	20.75 ± 13.03 (n=61)	.868	Hedges g = -0.02

Note. *Continuous variables are presented as mean ± standard deviation and were compared using the Mann-Whitney U test. Sex distribution was compared using the Yates-corrected chi-square test. Positive Hedges g values indicate a higher mean in the lithium group. AST = aspartate aminotransferase; ALT = alanine aminotransferase; BUN = blood urea nitrogen.*

Table 2. Data Structure, Follow-up Duration, and Outcome-Specific Analytic Samples

Measure	Value
Total patients	96
Treatment groups	Lithium: 35; Valproate: 61
Raw data rows / repeated records	282
Records per patient, mean ± SD	2.94 ± 0.50
Records per patient, median (range)	3 (1-4)
Follow-up, days, mean ± SD	312.22 ± 199.21
Follow-up, days, median (range)	322 (0-1105)
Creatinine: raw available records; adjusted model	279; 276 records / 95 patients
BUN: raw available records; adjusted model	280; 277 records / 95 patients
AST: raw available records; adjusted model	281; 278 records / 95 patients
ALT: raw available records; adjusted model	278; 275 records / 95 patients

Note. *Time was defined in months by dividing the number of days since each patient's first available laboratory record by 30. One patient with missing baseline age was excluded from the adjusted models. Because of outcome-specific missingness, the models were based on different numbers of observations.*

The raw dataset comprised 282 dated laboratory records from 96 patients. The mean number of usable records per patient was 2.94 ± 0.50, and the median was 3 (range, 1-4). This analytic record count is distinct from the eligibility requirement of at least three clinical follow-up visits because laboratory information was incomplete across visits and analytes. Mean observed laboratory follow-up was 312.22 ± 199.21 days, with a median of 322 days and a range of 0-1105 days. Because of missing outcome values and baseline age, adjusted models were based on 275-278 observations from 95 patients (Table 2).

Renal Outcomes

In the creatinine model, the time × treatment interaction was not statistically significant (B=-0.0019, 95% CI [-0.0056, 0.0017], p=.303); thus, the findings did not support a difference in creatinine slopes between treatment groups over the observed follow-up. The main effect of time was also not statistically significant (B=0.0029, p=.059). At the first available laboratory measurement, the adjusted creatinine level in the valproate group was nominally lower than that in the lithium group (B=-0.0659, 95% CI [-0.1314, -0.0003], p=.049). Because this coefficient represents an adjusted level difference at the start of observed laboratory follow-up, it was not interpreted as a longitudinal treatment effect. Male sex was associated with higher creatinine levels (B=0.1291, p<.001).

In the BUN model, the time × treatment interaction (B=-0.0853, 95% CI [-0.2042, 0.0337], p=.160), time (p=.286), and sex (p=.321) were not statistically significant. Older baseline age was associated with higher BUN levels (B=0.1262, 95% CI [0.0800, 0.1723], p<.001). Details of the renal models are presented in Table 3.

Table 3. Linear Mixed-Effects Model Findings for Renal Outcomes

Outcome	Predictor	B	SE	95% CI	z	p
Creatinine (276 records, 95 patients)	Time (months)	0.0029	0.0016	[-0.0001, 0.0060]	1.89	.059
	Valproate	-0.0659	0.0335	[-0.1314, -0.0003]	-1.97	.049
	Baseline age	0.0019	0.0012	[-0.0004, 0.0042]	1.65	.099
	Male	0.1291	0.0311	[0.0681, 0.1901]	4.15	<.001
	Time × valproate	-0.0019	0.0019	[-0.0056, 0.0017]	-1.03	.303
BUN (277 records, 95 patients)	Time (months)	0.0541	0.0507	[-0.0453, 0.1534]	1.07	.286
	Valproate	1.3553	0.7121	[-0.0403, 2.7509]	1.90	.057
	Baseline age	0.1262	0.0235	[0.0800, 0.1723]	5.36	<.001
	Male	0.6192	0.6236	[-0.6029, 1.8414]	0.99	.321
	Time × valproate	-0.0853	0.0607	[-0.2042, 0.0337]	-1.41	.160

Note. *B = unstandardized coefficient; SE = standard error; CI = confidence interval. Models were estimated using restricted maximum likelihood and included a patient-specific random intercept. Age was centered around the overall mean. Lithium and female sex were the reference categories. The valproate coefficient represents the adjusted level difference at observed time zero, whereas the time × valproate coefficient represents the monthly slope difference. The unconditional intraclass correlation coefficients for the creatinine and BUN models were 0.76 and 0.48, respectively.*

Table 4. Linear Mixed-Effects Model Findings for Hepatic Outcomes

Outcome	Predictor	B	SE	95% CI	z	p
log(AST+1) (278 records, 95 patients)	Time (months)	0.0062	0.0055	[-0.0044, 0.0169]	1.14	.252
	Valproate	-0.0573	0.0803	[-0.2147, 0.1002]	-0.71	.476
	Baseline age	-0.0017	0.0027	[-0.0069, 0.0035]	-0.63	.526
	Male	0.2530	0.0708	[0.1142, 0.3919]	3.57	<.001
	Time × valproate	-0.0070	0.0065	[-0.0197, 0.0058]	-1.07	.285
log(ALT+1) (275 records, 95 patients)	Time (months)	0.0042	0.0065	[-0.0085, 0.0169]	0.65	.516
	Valproate	-0.1212	0.1135	[-0.3436, 0.1012]	-1.07	.285
	Baseline age	-0.0002	0.0039	[-0.0078, 0.0074]	-0.06	.951
	Male	0.3690	0.1030	[0.1671, 0.5710]	3.58	<.001
	Time × valproate	-0.0059	0.0076	[-0.0209, 0.0091]	-0.77	.442

Note. *A log(x+1) transformation was applied to AST and ALT. With other variables held constant, the male coefficients correspond to approximately 28.8% and 44.6% higher geometric means for AST and ALT, respectively; these associations should not be interpreted as treatment toxicity. The unconditional intraclass correlation coefficients were 0.53 for AST and 0.67 for ALT. Other abbreviations and reference categories are defined in Table 3.*

Hepatic Outcomes

In the log-transformed AST model, the time × treatment interaction was not statistically significant (B=-0.0070, 95% CI [-0.0197, 0.0058], p=.285). Time, treatment group, and baseline age were also not associated with AST. Male sex was associated with higher AST levels (B=0.2530, p<.001); on the log scale, this coefficient corresponded to an approximately 28.8% higher geometric mean.

The time × treatment interaction was also not statistically significant in the ALT model (B=-0.0059, 95% CI [-0.0209, 0.0091], p=.442), and time, treatment group, and baseline age were not statistically significant. Male sex was associated with higher ALT levels (B=0.3690, p<.001), corresponding to an approximately 44.6% higher geometric mean. Details of the hepatic models are presented in Table 4.

Secondary and Sensitivity Analyses

In the piecewise creatinine model, the slope during the first 6 months was B=0.0065 (p=.093) in the lithium group and B=0.0041 (p=.123) in the valproate group. Slopes after month 6 were not statistically significant in either group. More importantly for the comparative research question, the treatment-by-period interaction estimates did not support differential early or late trajectories: early-period interaction B=-0.0024 (95% CI [-0.0116, 0.0069], p=.614) and late-period interaction B=-0.0005 (95% CI [-0.0079, 0.0069], p=.888); the joint Wald test was $\chi^2(2)=0.544$, p=.762. The findings therefore did not support treatment-specific early or late creatinine change.

Across the random-slope models and sensitivity analyses excluding repeated panel records, patients followed for more than 18 months, or patients with fewer than three analytic records, the time × treatment interaction did not become statistically significant for any of the four outcomes. When all data-quality restrictions were applied simultaneously, interaction p values were .445 for creatinine, .291 for BUN, .199 for AST, and .365 for ALT (Table 5).

Table 5. Piecewise Time Model and Sensitivity Analyses for Creatinine

Panel A. Treatment-Specific Piecewise Model With a 6-Month Knot				
Estimate / test	**B or χ^2**	**SE**	**95% CI**	**p**
Early slope: lithium	0.0065	0.0039	[-0.0011, 0.0141]	.093
Early slope: valproate	0.0041	0.0027	[-0.0011, 0.0094]	.123
Late slope: lithium	-0.0001	0.0034	[-0.0068, 0.0066]	.977
Late slope: valproate	-0.0006	0.0017	[-0.0039, 0.0026]	.704
Early slope × treatment difference	-0.0024	0.0047	[-0.0116, 0.0069]	.614
Late slope × treatment difference	-0.0005	0.0038	[-0.0079, 0.0069]	.888
Joint Wald test of treatment interactions	$\chi^2(2)=0.544$	—	—	.762
Note. *The model was based on 276 records from 95 patients; it was adjusted for baseline age, sex, and treatment group and included a patient-specific random intercept. Early and late time were defined as the first 6 months and the period after month 6, respectively, following the first available laboratory measurement.*				
Panel B. Sensitivity Analysis p Values for the Time × Treatment Interaction				
Analysis	Creatinine	BUN	AST	ALT
Primary random-intercept model	.303	.160	.285	.442
Random-slope model	.200	.224	.183	.362
Later record from identical panel excluded	.332	.153	.277	.448
Four patients with follow-up >18 months excluded	.520	.474	.166	.257
Patients with at least three analytic records	.273	.147	.352	.588
All data-quality restrictions applied simultaneously	.445	.291	.199	.365
Note. *The time × treatment interactions remained statistically nonsignificant in all sensitivity analyses. Because slope variances for creatinine, AST, and ALT were at the boundary in the random-slope models, these results were interpreted as supportive sensitivity findings.*				

Discussion

This study examined within-person trajectories of creatinine, BUN, AST, and ALT using repeated clinical measurements from patients with bipolar disorder treated with lithium or valproate. In models adjusted for baseline age and sex, the time × treatment interaction was not statistically significant for any of the four markers. The piecewise analysis of treatment-specific early and late creatinine slopes and multiple data-quality sensitivity analyses did not alter this primary finding. These results indicate that average biochemical slopes were not statistically distinguishable between treatment groups over the observed follow-up period; they do not provide evidence of equivalence, absence of individual toxicity, or long-term safety.

The magnitude and precision of the interaction estimates are also important. For creatinine, the monthly treatment-by-time difference was B=-0.0019 with a 95% CI from -0.0056 to 0.0017 mg/dL/month; for BUN it was B=-0.0853 with a 95% CI from -0.2042 to 0.0337 mg/dL/month. For

log(AST+1) and log(ALT+1), the corresponding interaction estimates were B=-0.0070 (95% CI [-0.0197, 0.0058]) and B=-0.0059 (95% CI [-0.0209, 0.0091]), respectively. These intervals include zero but also encompass modest effects in either direction, so the results support absence of a statistically detectable difference rather than precise exclusion of clinically meaningful small differences.

Renal risk associated with lithium may vary according to exposure duration and the outcome examined. Although meta-analytic evidence suggests that the average decline in glomerular filtration rate may be limited, studies with longer follow-up have indicated that chronic loss of renal function may become more apparent with older age, higher serum lithium concentrations, and longer exposure (15-19). In the present study, the absence of statistically detectable differences in creatinine and BUN slopes relative to valproate may be related to the mean follow-up of approximately 10 months, the exclusion of patients with chronic renal failure at baseline, and the lack of more sensitive outcomes such as estimated glomerular filtration rate (eGFR) and albuminuria. The absence of change in the creatinine slope does not exclude impaired urinary concentrating ability or small but clinically relevant losses of eGFR at the individual level.

Hepatic effects associated with valproate range from transient aminotransferase elevations to rare but severe idiosyncratic liver injury (21-23). Hyperammonemia may also occur despite normal AST and ALT levels (24). The absence of marked divergence in mean aminotransferase changes in clinical studies directly comparing lithium and valproate is consistent with the present findings (25,26). Nevertheless, the average slope of continuous laboratory values may obscure elevations beyond clinical thresholds, events requiring treatment discontinuation, or rare idiosyncratic toxicity. The lack of a difference in AST and ALT slopes therefore does not indicate that valproate is free of hepatic risk.

The associations of male sex with creatinine, AST, and ALT levels do not appear specific to treatment toxicity. Differences in muscle mass underlying creatinine production and sex-related variation in renal physiology may contribute to higher creatinine levels (20). Similarly, body composition, metabolic risk, alcohol use, and lifestyle factors may influence the distribution of liver enzyme levels by sex (27,28). The limited detail available for these variables in the source records constrains mechanistic interpretation.

The piecewise analysis compared early and late periods after the first available laboratory record and did not identify a treatment-specific early creatinine slope. The emphasis is therefore on the treatment-by-period interactions rather than on any unadjusted common early slope. This 6-month window is clinically relevant because contemporary guidance for lithium treatment recommends close early therapeutic drug monitoring and continued surveillance of renal function, including eGFR, during maintenance treatment (30). Our result should not be interpreted as evidence that intensive early monitoring is unnecessary; rather, no differential mean creatinine slope between lithium and valproate was detectable within the observed early window. Because severe hepatic reactions associated with valproate may also occur early, the present laboratory-only findings should not be used to modify established valproate monitoring recommendations (29).

Strengths of this study include the use of mixed-effects models that accounted for within-person dependence in irregularly spaced real-world measurements, coding age as a patient-level baseline value, transparent reporting of the number of observations in each model, and evaluation of the findings across different data-quality scenarios. The consistency of the primary result in random-slope models provided additional support for the findings from the primary random-intercept models.

These findings should be interpreted in light of several limitations. The retrospective, single-center design limits causal inference and generalizability. Treatment assignment was not randomized, and age and sex differed between the lithium and valproate groups, indicating potential confounding by indication. Statistical adjustment can only partially address this issue. Serum lithium and valproate concentrations, prescribed dose, dose changes, adherence, bipolar disorder subtype, episode characteristics, previous treatment response, concomitant psychotropic treatment (including second-generation antipsychotics and other agents with potential metabolic or hepatic effects), and relevant medical comorbidities were not available with sufficient consistency for modeling. Consequently, subtherapeutic exposure, dose fluctuation, non-adherence, or polypharmacy could have attenuated or otherwise altered observed treatment-by-time differences. The study should therefore be interpreted as describing average biochemical trajectories under naturalistic clinical observation rather than isolating controlled drug-specific toxicity. Selection of patients with regular follow-up and

available laboratory records may also have preferentially retained individuals who continued treatment and monitoring; patients who discontinued treatment or were lost to follow-up because of adverse effects could have been underrepresented, introducing selection bias.

Because criterion-specific exclusion counts were not retained in the analytic file, the transition from 664 screened records to 96 included patients could not be partitioned by individual exclusion reason. Although eligibility required at least three clinical follow-up visits, laboratory measurements were incomplete across visits and analytes; consequently, individual patients contributed only one to four usable laboratory records. This incompleteness limits the characterization of within-person biochemical trajectories. The raw data contained four follow-up periods exceeding 18 months and seven apparently identical laboratory-panel pairs recorded 7-8 days apart. Although sensitivity analyses excluding these records did not alter the findings, the inability to verify them in the source system represents a data-quality limitation. The time axis begins at the first available laboratory measurement and does not represent time since treatment initiation; early treatment-related changes occurring before that measurement may therefore have been missed. Creatinine and BUN were the renal markers consistently available in the analytic dataset, but they are incomplete proxies for renal safety. Serum creatinine is influenced by age, sex, muscle mass, hydration, and other factors, while BUN is likewise nonspecific. Contemporary lithium-monitoring guidance incorporates eGFR because it provides a more clinically interpretable estimate of filtration (30). Laboratory-reported eGFR was not available in the source analytic dataset, and a validated visit-level eGFR series could not be independently verified retrospectively; this limits renal inference and is particularly relevant given the observed age and sex imbalance between treatment groups. Missing laboratory outcomes were assumed to be missing at random conditional on observed variables. With 96 patients and 275-278 outcome-specific observations, the repeated-measures design provides more information than a single cross-sectional measurement per patient, but the effective precision of treatment-by-time interactions also depends on within-person correlation, irregular observation timing, missingness, and slope heterogeneity. The confidence intervals around the primary interactions remain compatible with small effects in either direction; therefore, Type II error cannot be excluded and the study is not sufficiently precise to rule out small clinically meaningful between-treatment differences. Finally, creatinine and BUN do not comprehensively represent renal safety, nor do AST and ALT comprehensively represent hepatic safety; eGFR, albuminuria, urinary concentrating capacity, bilirubin, international normalized ratio, albumin, ammonia, and clinical adverse events were unavailable.

Conclusion

No statistically detectable difference was observed between lithium and valproate groups in the longitudinal slopes of creatinine, BUN, AST, and ALT over the observed follow-up. Treatment-specific early or late creatinine slopes were also not supported, and the findings were consistent across different data-quality scenarios. These results do not demonstrate equivalence or long-term comparative safety. The estimates were imprecise, and the available renal and hepatic marker set was incomplete. Accordingly, these findings should not be used to reduce established drug-specific monitoring. Larger prospective studies with longer follow-up should jointly assess serum drug concentrations, dose, adherence, eGFR, albuminuria, metabolic risks, concomitant medications, and clinical adverse events.

Declarations

Funding

No financial support was received from any institution, organization, or commercial source for this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest related to this study.

Ethics Approval

The study was approved by the Non-Interventional Scientific Research Ethics Committee of Tokat Gaziosmanpaşa University Faculty of Medicine (Decision No: 25-MOBAEK-355; Date: 14.10.2025). Individual informed consent was not obtained because the study was based on a retrospective review of medical records.

Availability of Data and Material

The de-identified data used in this study are available from the corresponding author upon reasonable request,

subject to ethical and institutional restrictions and provision of an appropriate justification.

Authors' Contributions

İG: conceptualization and study design, methodology, statistical analysis, interpretation of findings, preparation of the initial draft, and critical revision of the manuscript. SA: study design, data collection, data validation, clinical interpretation, and critical revision of the manuscript. BE: data collection, data curation, literature review, clinical interpretation of the findings, and critical revision of the manuscript. All authors read and approved the final version of the manuscript and agreed to be accountable for the accuracy and integrity of the work.

Acknowledgment

Not applicable.

References

1. Oliva V, Fico G, De Prisco M, et al. Bipolar disorders: an update on critical aspects. Lancet Reg Health Eur. 2025;48:101135.
2. Geddes JR, Miklowitz DJ. Treatment of bipolar disorder. Lancet. 2013;381:1672-82.
3. Nierenberg AA, Agustini B, Köhler-Forsberg O, et al. Diagnosis and treatment of bipolar disorder: a review. JAMA. 2023;330:1370-80.
4. Tondo L, Abramowicz M, Alda M, et al. Long-term lithium treatment in bipolar disorder: effects on glomerular filtration rate and other metabolic parameters. Int J Bipolar Disord. 2017;5:27.
5. Malhi GS, Gessler D, Outhred T. The use of lithium for the treatment of bipolar disorder: recommendations from clinical practice guidelines. J Affect Disord. 2017;217:266-80.
6. Airainer M, Seifert R. Lithium, the gold standard drug for bipolar disorder: analysis of current clinical studies. Naunyn Schmiedebergs Arch Pharmacol. 2024;397:9723-43.
7. Yee CS, Vazquez GH, Hawken ER, et al. Long-term treatment of bipolar disorder with valproate: updated systematic review and meta-analyses. Harv Rev Psychiatry. 2021;29:188-95.
8. Hayes JF, Marston L, Walters K, et al. Adverse renal, endocrine, hepatic, and metabolic events during maintenance mood stabilizer treatment for bipolar disorder: a population-based cohort study. PLoS Med. 2016;13:e1002058.
9. Chan JKN, Solmi M, Correll CU, et al. Lithium for bipolar disorder and risk of thyroid dysfunction and chronic kidney disease. JAMA Netw Open. 2025;8:e2458608.
10. Bosi A, Clase CM, Ceriani L, et al. Absolute and relative risks of kidney outcomes associated with lithium vs valproate use in Sweden. JAMA Netw Open. 2023;6:e2322056.
11. Rej S, Herrmann N, Shulman K, et al. Lithium use, but not valproate use, is associated with a higher risk of chronic kidney disease in older adults with mental illness. J Clin Psychiatry. 2017;78:e980-5.
12. Kessing LV, Knudsen MB, Rytgaard HCW, et al. Lithium versus anticonvulsants and the risk of physical disorders: results from a comprehensive long-term nationwide population-based study emulating a target trial. Eur Neuropsychopharmacol. 2024;84:48-56.
13. Wiuff AC, Rohde C, Jensen BD, et al. Association between lithium treatment and renal, thyroid and parathyroid function: a cohort study of 6659 patients with bipolar disorder. Bipolar Disord. 2024;26:71-83.
14. Mannapperuma U, Galappatthy P, Jayakody RL, et al. Safety monitoring of treatment in bipolar disorder in a tertiary care setting in Sri Lanka and recommendations for improved monitoring in resource-limited settings. BMC Psychiatry. 2019;19:194.
15. McKnight RF, Adida M, Budge K, et al. Lithium toxicity profile: a systematic review and meta-analysis. Lancet. 2012;379:721-8.
16. Turan T, Eşel E, Tokgöz B, et al. Effects of short- and long-term lithium treatment on kidney functioning in patients with bipolar mood disorder. Prog Neuropsychopharmacol Biol Psychiatry. 2002;26:561-5.
17. Shine B, McKnight RF, Leaver L, et al. Long-term effects of lithium on renal, thyroid, and parathyroid function: a retrospective analysis of laboratory data. Lancet. 2015;386:461-8.
18. Rej S, Herrmann N, Gruneir A, et al. Association of lithium use and a higher serum concentration of lithium with the risk of declining renal function in older adults: a population-based cohort study. J Clin Psychiatry. 2020;81:19m13045.
19. Cho Y, Lee D, Baek JH, et al. Estimated glomerular filtration rate in Korean patients exposed to long-term lithium maintenance therapy. Int J Bipolar Disord. 2022;10:4.
20. Harvey BJ, Alvarez De La Rosa D. Sex differences in kidney health and disease. Nephron. 2025;149:77-103.
21. Zaccara G, Franciotta D, Perucca E. Idiosyncratic adverse reactions to antiepileptic drugs. Epilepsia. 2007;48:1223-44.
22. Kadam R, Palkar M, Pingili RB. Mechanisms involved in valproic acid-induced hepatotoxicity: a comprehensive review. Toxicol Mech Methods. 2025;35:565-80.
23. Guo HL, Jing X, Sun JY, et al. Valproic acid and liver injury in patients with epilepsy: an update. Curr Pharm Des. 2019;25:343-51.
24. Wadzinski J, Franks D, Roane D, et al. Valproate-associated hyperammonemic encephalopathy. J Am Board Fam Med. 2007;20:499-502.
25. Bowden CL, Mosolov S, Hranov L, et al. Efficacy of valproate versus lithium in mania or mixed mania: a randomized, open 12-week trial. Int Clin Psychopharmacol. 2010;25:60-7.
26. Kang MG, Qian H, Keramatian K, et al. Lithium vs valproate in the maintenance treatment of bipolar I disorder: a post-hoc analysis of a randomized double-blind placebo-controlled trial. Aust N Z J Psychiatry. 2020;54:298-307.
27. Niemelä O, Bloigu A, Bloigu R, et al. Associations between liver enzymes, lifestyle risk factors and pre-existing medical conditions in a population-based cross-sectional sample. J Clin Med. 2023;12:4276.
28. Qiu J, Kuang M, He S, et al. Gender perspective on the association between liver enzyme markers and non-alcoholic fatty liver disease: insights from the general population. Front Endocrinol. 2023;14:1302322.
29. National Institute of Diabetes and Digestive and Kidney Diseases. LiverTox: clinical and research information on drug-induced liver injury. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases; 2012.
30. National Institute for Health and Care Excellence. Bipolar disorder: assessment and management. Clinical guideline CG185. Updated 2023. Available from: https://www.nice.org.uk/guidance/cg185. Accessed September 2, 2026.