Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Comprehensive Analysis of Liver Enzyme Profiles: Age-Related Variations, Predictive Modeling, and Phenotypic Classification

Dr. Vikas Tiwari

vikas.tiwari.vats@gmail.com

AIPH University

Dr. Jaishree Tiwari

AIPH University

Research Article

Keywords: Liver enzymes, AST, ALT, ALP, regression analysis, cluster analysis, ANOVA, biomarkers

Posted Date: April 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9263184/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background: Liver enzymes are critical biomarkers reflecting hepatocellular and cholestatic injury. However, their variability across demographic factors and their interrelationships remain incompletely understood.

Objective: To evaluate the distribution, determinants, and clustering patterns of liver enzymes (AST, ALT, ALP) using advanced statistical approaches.

Methods: A cross-sectional study was conducted on 50 participants. Descriptive statistics, independent t-tests, Pearson correlation, one-way ANOVA, and multiple linear regression were performed. Log transformation was applied to normalize skewed data. Outlier diagnostics and K-means clustering were used for advanced analysis.

Results: The mean age was 21.42 ± 11.02 years. No significant gender differences were observed ($p > 0.05$). ALP showed significant variation across age groups ($p = 0.015$). Regression analysis identified age ($\beta = -1.12$, $p = 0.018$) and AST ($\beta = 0.82$, $p = 0.006$) as independent predictors of ALP. Log transformation improved model stability. Cluster analysis identified three distinct biochemical phenotypes: normal, hepatocellular, and cholestatic patterns.

Conclusion: ALP is the most variable and clinically informative enzyme, significantly influenced by age and AST. Advanced statistical approaches reveal distinct biochemical phenotypes, supporting heterogeneity in liver function dynamics.

INTRODUCTION

Liver function tests (LFTs) are among the most commonly utilized biochemical investigations in clinical practice, serving as essential tools for the evaluation of hepatic integrity, metabolic function, and systemic health. Among these, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and alkaline phosphatase (ALP) are key enzymes that provide valuable insights into hepatocellular injury and cholestatic processes. Despite their widespread use, interpretation of these biomarkers remains complex due to physiological variability, demographic influences, and overlapping pathological mechanisms.

AST and ALT are intracellular enzymes involved in amino acid metabolism and are primarily released into the circulation following hepatocellular damage. ALT is considered more liver-specific, whereas AST is also present in extrahepatic tissues such as cardiac and skeletal muscle, limiting its specificity[1]. In contrast, ALP is predominantly associated with the biliary epithelium and is widely regarded as a marker of cholestasis, although it is also produced in bone and other tissues[2]. This diversity of sources complicates clinical interpretation, particularly in cases where multiple enzymes are concurrently elevated.

The interpretation of liver enzyme levels is further influenced by demographic factors such as age and gender. Several studies have demonstrated that aminotransferase levels tend to decline with advancing age, possibly due to reduced hepatic mass, diminished enzymatic activity, or changes in metabolic demand[3]. Conversely, younger individuals may exhibit relatively higher enzyme levels, which may reflect increased metabolic activity or transient subclinical inflammation[4]. Gender-related differences have also been reported, with males often showing higher ALT levels, potentially due to hormonal influences, body composition, and lifestyle factors[5]. However, these findings are not consistent across all populations, highlighting the need for context-specific analysis.

In addition to demographic variability, liver enzyme levels are influenced by a wide range of physiological and pathological conditions, including metabolic syndrome, obesity, alcohol consumption, infections, and drug-induced liver injury[6]. Mild elevations in liver enzymes, even within the upper limit of normal, have been associated with increased risk of cardiovascular disease and metabolic disorders, underscoring their role as early indicators of systemic dysfunction[7]. Consequently, there is growing interest in exploring the broader clinical significance of these biomarkers beyond traditional liver diseases.

Conventional approaches to analyzing liver enzyme data often rely on descriptive statistics and simple comparisons, which may not adequately capture the complex relationships between variables. In recent years, advanced statistical techniques have been increasingly applied to biomedical research to address these limitations. Methods such as analysis of variance (ANOVA), multivariate regression, and cluster analysis enable a more comprehensive understanding of data patterns and underlying associations[8].

One of the major challenges in biochemical data analysis is the presence of skewed distributions, particularly for enzymes such as ALP, which can exhibit substantial variability and extreme values. Skewness can violate the assumptions of parametric statistical tests, leading to biased estimates and reduced validity. Logarithmic transformation is commonly employed to normalize such data, improving statistical robustness and interpretability[9]. Proper handling of data distribution is therefore essential for accurate analysis and reliable conclusions.

Another critical aspect of statistical analysis is the identification of outliers and influential observations. In clinical datasets, extreme values may represent true pathological conditions or measurement errors. Techniques such as standardized residuals and Cook’s distance are widely used to detect such observations and assess their impact on statistical models[10]. Careful evaluation of outliers is necessary to ensure that meaningful clinical information is not inadvertently excluded.

Beyond traditional statistical methods, cluster analysis has emerged as a powerful tool for identifying hidden patterns within complex datasets. By grouping individuals based on similarities in their biochemical profiles, cluster analysis can reveal distinct phenotypic subgroups that may correspond to different pathophysiological mechanisms[11]. In the context of liver enzymes, this approach can help differentiate between hepatocellular and cholestatic patterns, providing additional diagnostic and prognostic insights.

Furthermore, multivariate regression analysis allows for the identification of independent predictors of enzyme levels, taking into account the simultaneous effects of multiple variables. This is particularly important in understanding the interplay between different liver enzymes and demographic factors. For instance, the relationship between AST and ALP may reflect overlapping pathways of hepatic injury and biliary dysfunction, which cannot be fully appreciated through univariate analysis alone[12].

The integration of these advanced analytical techniques represents a significant advancement in the study of liver enzyme dynamics. By moving beyond simple descriptive approaches, researchers can gain deeper insights into the factors influencing enzyme variability and their clinical implications. This is particularly relevant in the era of precision medicine, where individualized assessment and targeted interventions are increasingly emphasized[13].

Despite these advancements, there remains a need for comprehensive studies that incorporate multiple statistical approaches to evaluate liver enzyme patterns in a unified framework. Many existing studies are limited by small sample sizes, lack of data normalization, or insufficient exploration of multivariate relationships[14]. Addressing these gaps is essential for improving the accuracy and clinical relevance of liver enzyme interpretation.

In this context, the present study aims to provide a detailed analysis of liver enzyme profiles using a combination of descriptive, inferential, and advanced statistical methods. By examining the influence of age and gender, assessing variability through ANOVA, identifying predictors □□□□□□□ regression modeling, and exploring phenotypic patterns □□□□□□□ cluster analysis, this study seeks to offer a comprehensive understanding of liver enzyme dynamics. Such an approach not only enhances statistical rigor but also contributes to more informed clinical decision-making.

METHODOLOGY

Study Design and Population

A cross-sectional analytical study was conducted on a dataset comprising 50 participants. The study included individuals across a broad age range, allowing evaluation of demographic influences on liver enzyme profiles. Participants with available data on age, gender, and biochemical parameters were included in the analysis. No exclusion was applied based on enzyme values to preserve the natural variability of the dataset.

Data Collection and Variables

The dataset included demographic variables (age and gender) and biochemical parameters, namely aspartate aminotransferase (AST), alanine aminotransferase (ALT), and alkaline phosphatase (ALP). All enzyme values were recorded in standard international units per liter (IU/L). Age was treated as a

continuous variable and also categorized into groups (≤ 20, 21–40, 41–60, > 60 years) for comparative analysis.

Statistical Analysis

Statistical analysis was performed using standard analytical approaches. Continuous variables were expressed as mean ± standard deviation (SD), while categorical variables were summarized as frequencies and percentages.

Comparative Analysis

- **Independent sample t-test** was used to compare mean enzyme levels between gender groups.
- **One-way analysis of variance (ANOVA)** was applied to assess differences in enzyme levels across age groups.

Correlation and Regression Analysis

- **Pearson correlation coefficient** was used to evaluate the relationship between age and liver enzyme levels.
- **Multiple linear regression analysis** was performed to identify independent predictors of ALP, with age, AST, ALT, and gender included as explanatory variables.

Data Transformation and Assumption Testing

Given the skewed distribution of ALP, a logarithmic transformation was applied to improve normality and meet the assumptions of parametric tests. Normality of variables was assessed using graphical methods and standard statistical tests.

Outlier and Influence Diagnostics

Outliers and influential observations were evaluated using standardized residuals and Cook's distance. Observations exceeding accepted thresholds were examined for their impact on regression models. Sensitivity analysis was performed to ensure robustness of findings.

Cluster Analysis

To identify underlying biochemical patterns, K-means cluster analysis was performed using AST, ALT, and ALP values. The optimal number of clusters was determined based on interpretability and variance distribution, resulting in classification into distinct phenotypic groups.

Statistical Significance

A **p-value < 0.05** was considered statistically significant for all analyses.

RESULTS

Baseline Characteristics

A total of 50 participants were included in the present study, representing a relatively young population with a mean age of 21.42 ± 11.02 years. The wide standard deviation indicates substantial age dispersion within the cohort, allowing for meaningful evaluation of age-related variability in liver enzyme parameters. The inclusion of individuals across different age brackets enhances the analytical robustness of comparisons and trend analyses.

The overall biochemical profile demonstrated the following mean enzyme levels:

- **Aspartate aminotransferase (AST)**: 29.02 ± 12.90 IU/L
- **Alanine aminotransferase (ALT)**: 26.56 ± 13.92 IU/L
- **Alkaline phosphatase (ALP)**: 150.10 ± 115.21 IU/L

Among the three enzymes analyzed, ALP exhibited the highest degree of variability, as evidenced by its markedly elevated standard deviation relative to its mean. This suggests the presence of substantial dispersion and possible extreme values within the dataset. In contrast, AST and ALT demonstrated relatively moderate variability, indicating a more homogeneous distribution across the study population.

The distribution pattern of AST and ALT appeared relatively symmetric, with values clustered around the mean, consistent with near-normal distribution characteristics. However, ALP demonstrated a positively skewed distribution, indicating the presence of higher-end outliers or subgroups with markedly elevated values. This skewness is clinically relevant, as it may reflect underlying heterogeneity in cholestatic activity or metabolic processes within the cohort.

The observed baseline enzyme levels fall within generally accepted physiological ranges; however, the wide variability in ALP suggests that certain individuals may exhibit subclinical or clinically significant elevations. This heterogeneity underscores the importance of employing advanced statistical techniques to further explore underlying patterns and associations.

Gender-wise Comparison

The study cohort included both male and female participants, enabling comparative analysis of liver enzyme levels across gender groups. Independent sample t-tests were conducted to evaluate potential differences in mean enzyme values between males and females.

The analysis revealed no statistically significant differences in AST, ALT, or ALP levels between the two groups ($p > 0.05$ for all comparisons). This finding suggests that, within this study population, liver enzyme levels are not significantly influenced by gender.

Although minor variations in mean values were observed between males and females, these differences did not reach statistical significance and are likely attributable to random variation rather than true biological differences. This observation is important, as it indicates that gender may not be a major determinant of liver enzyme variability in relatively homogeneous or young populations.

The lack of significant gender-based differences also supports the validity of combining male and female data for subsequent analyses, thereby enhancing statistical power. Furthermore, this finding aligns with emerging evidence suggesting that gender differences in liver enzymes are often context-dependent and may be influenced by confounding factors such as lifestyle, hormonal status, and metabolic conditions rather than intrinsic biological variation.

Correlation Analysis

Pearson correlation analysis was performed to assess the relationship between age and liver enzyme levels. The results demonstrated weak negative correlations between age and all three enzymes (AST, ALT, and ALP). However, these associations did not reach statistical significance ($p > 0.05$).

The negative direction of the correlations suggests a general trend of decreasing enzyme levels with increasing age. Although this trend was not statistically significant, it may indicate a subtle age-related decline in hepatic enzyme activity or metabolic turnover.

The weak strength of the correlations implies that age alone is not a strong predictor of enzyme levels in this dataset. This finding highlights the multifactorial nature of liver enzyme regulation, where multiple physiological and environmental factors contribute to observed variability.

It is also noteworthy that the absence of statistical significance may be influenced by the relatively small sample size, which limits the power to detect subtle associations. Nevertheless, the observed trend is consistent with previously reported patterns in the literature, suggesting that enzyme levels may decline with age due to reduced hepatic mass or metabolic activity.

One-Way ANOVA

To further explore age-related differences, participants were categorized into predefined age groups, and one-way analysis of variance (ANOVA) was performed.

The analysis revealed that:

- ALP levels differed significantly across age groups ($p = 0.015$)
- AST and ALT did not show statistically significant differences ($p > 0.05$)

The significant variation in ALP across age groups indicates that age has a measurable impact on ALP levels, even in the absence of strong linear correlation. This suggests that the relationship between age and ALP may be non-linear or influenced by specific age-related physiological changes.

Post-hoc observations (trend-based) indicated that younger individuals tended to have higher ALP levels, while older individuals exhibited relatively lower values. This pattern may reflect differences in bone metabolism, growth-related processes, or biliary function.

In contrast, the absence of significant differences for AST and ALT suggests that these enzymes are relatively stable across age groups within this population. This stability may indicate that aminotransferase levels are less sensitive to age-related changes compared to ALP.

Overall, the ANOVA findings highlight the distinct behavior of ALP compared to AST and ALT, reinforcing its role as a variable and potentially informative biomarker.

Multiple Linear Regression

To identify independent predictors of ALP levels, a multiple linear regression analysis was performed, incorporating age, AST, ALT, and gender as explanatory variables.

The regression model demonstrated that:

- Age was a significant negative predictor of ALP ($p = 0.018$)
- AST was a significant positive predictor of ALP ($p = 0.006$)
- ALT and gender were not statistically significant predictors

The model explained approximately 31% of the variance in ALP levels ($R^2 = 0.31$), indicating a moderate level of explanatory power.

The negative association between age and ALP confirms the trend observed in earlier analyses, suggesting that ALP levels tend to decline with increasing age. This relationship may be influenced by age-related changes in bone turnover or biliary activity.

The positive association between AST and ALP suggests a potential interaction between hepatocellular and cholestatic processes. Elevated AST levels may reflect hepatocellular injury, which could indirectly influence ALP levels through shared or overlapping pathological pathways.

The lack of significance for ALT indicates that it may not play a major role in determining ALP variability in this dataset. This finding is consistent with the notion that ALT is more specific to hepatocellular injury and may not directly correlate with cholestatic markers.

Overall, the regression analysis provides important insights into the independent determinants of ALP, highlighting the combined influence of age and AST.

Log Transformation

Given the skewed distribution of ALP, a logarithmic transformation was applied to normalize the data.

The transformation resulted in:

- Reduction in skewness
- Improved approximation to normal distribution
- Enhanced reliability of statistical analyses

Following transformation, the distribution of ALP became more symmetric, satisfying the assumptions required for parametric tests such as regression and ANOVA.

This step is particularly important, as skewed data can lead to biased estimates and inaccurate conclusions. By normalizing the distribution, log transformation ensures that statistical models provide more valid and interpretable results.

Outlier Diagnostics

Outlier and influence diagnostics were conducted using standardized residuals and Cook's distance.

The analysis identified a small number of influential observations, primarily associated with markedly elevated ALP values. These observations were carefully evaluated to determine their impact on the regression model.

Sensitivity analysis demonstrated that:

- Exclusion of these observations did not significantly alter the results
- Key associations (age and AST with ALP) remained consistent

This indicates that the findings are robust and not driven by extreme values. Importantly, these outliers likely represent clinically meaningful cases rather than errors, and their inclusion provides valuable insight into the full spectrum of biochemical variability.

Cluster Analysis

To explore underlying patterns in enzyme profiles, K-means cluster analysis was performed using AST, ALT, and ALP values.

The analysis identified three distinct clusters, representing different biochemical phenotypes:

Cluster 1: Normal Pattern

- Characterized by low or near-normal levels of all three enzymes
- Represents individuals with minimal biochemical abnormalities

Cluster 2: Hepatocellular Pattern

- Elevated AST and ALT levels
- Moderate ALP levels
- Suggestive of hepatocellular injury

Cluster 3: Cholestatic Pattern

- Markedly elevated ALP levels
- Relatively lower AST and ALT
- Indicative of cholestatic or biliary involvement

The identification of these clusters highlights the heterogeneity of liver enzyme profiles within the study population. This phenotypic classification provides a more nuanced understanding of biochemical patterns compared to traditional single-parameter analysis.

Cluster analysis also demonstrates the potential for data-driven stratification, which may have important clinical implications for diagnosis and management.

SUMMARY OF KEY FINDINGS

- ALP showed highest variability and significant age-related differences
- No gender-based differences observed
- Age and AST are independent predictors of ALP
- Log transformation improved statistical validity
- Results are robust despite outliers
- Cluster analysis revealed three distinct biochemical phenotypes

DISCUSSION

The present study provides a comprehensive evaluation of liver enzyme dynamics using both conventional and advanced statistical techniques, offering important insights into the variability, determinants, and interrelationships of AST, ALT, and ALP. The findings highlight the complex nature of liver enzyme regulation and underscore the importance of adopting multidimensional analytical approaches in clinical research.

One of the key observations of this study is the lack of significant gender-based differences in liver enzyme levels. This finding is consistent with several previous studies that have reported minimal or no differences in aminotransferase and ALP levels between males and females when confounding factors such as alcohol consumption and metabolic status are controlled[15]. Although some earlier reports suggested higher ALT levels in males, these differences are often attributed to lifestyle and hormonal influences rather than intrinsic biological variation[16]. The absence of significant gender differences in the present study suggests that liver enzyme variability may be more strongly influenced by other factors, such as age and metabolic activity.

Age-related variation in liver enzymes represents another important aspect of this study. Although correlation analysis demonstrated a weak inverse relationship between age and enzyme levels, only ALP showed statistically significant variation across age groups in the ANOVA model. This finding aligns with previous research indicating that ALP levels tend to decline with age, possibly due to reduced bone turnover and changes in biliary function[17]. The observed trend of higher ALP levels in younger individuals may reflect increased osteoblastic activity or transient physiological variations, particularly in growing individuals or those with higher metabolic demands[18].

The identification of age as a significant predictor of ALP in the regression model further reinforces its role in influencing enzyme variability. The negative regression coefficient indicates that ALP decreases with increasing age, a finding that is consistent with epidemiological studies demonstrating age-dependent changes in biochemical markers[19]. This highlights the importance of considering age-specific reference ranges when interpreting liver enzyme levels, as a single cutoff value may not be appropriate across all age groups.

Another notable finding of this study is the significant association between AST and ALP in the multivariate regression model. This relationship suggests a potential overlap between hepatocellular and cholestatic processes, supporting the concept that liver injury often involves multiple pathways rather than isolated mechanisms[20]. Elevated AST levels may reflect hepatocellular damage, which in turn can affect biliary function and lead to increased ALP levels. This interconnectedness underscores the need for integrated interpretation of liver enzymes rather than reliance on individual markers.

In contrast, ALT did not emerge as a significant predictor of ALP in the regression analysis. This may be due to its greater specificity for hepatocellular injury, which does not necessarily correlate with cholestatic processes. Previous studies have similarly reported weaker associations between ALT and ALP compared to AST, suggesting that these enzymes may reflect distinct aspects of liver function[21]. The differential behavior of AST and ALT highlights the importance of considering the unique physiological roles of each enzyme when interpreting results.

The application of log transformation to ALP represents a critical methodological strength of this study. Given the highly skewed distribution of ALP, transformation was necessary to meet the assumptions of parametric statistical tests and improve model accuracy. The successful normalization of ALP following transformation is consistent with established statistical principles and emphasizes the importance of data preprocessing in biomedical research[22]. Failure to address skewness can lead to misleading conclusions, particularly in regression analyses where outliers can exert disproportionate influence.

Outlier and influence diagnostics further demonstrated the robustness of the study findings. Although a small number of influential observations were identified, their exclusion did not significantly alter the results, indicating that the overall conclusions are stable and reliable. Importantly, these extreme values likely represent clinically relevant cases rather than measurement errors, highlighting the need to carefully evaluate outliers before deciding on their inclusion or exclusion[23]. This approach ensures that meaningful clinical variability is preserved while minimizing statistical bias.

The use of cluster analysis in this study provides valuable insights into the heterogeneity of liver enzyme profiles. The identification of three distinct clusters—normal, hepatocellular, and cholestatic patterns—supports the concept that liver enzyme abnormalities can be categorized into discrete phenotypic groups. This finding is consistent with previous studies that have utilized clustering techniques to identify subgroups within clinical populations[24]. Such phenotypic classification has important clinical implications, as it can guide diagnostic evaluation and inform treatment strategies.

The hepatocellular cluster, characterized by elevated AST and ALT, likely represents individuals with predominant hepatocellular injury, which may be associated with conditions such as viral hepatitis or drug-induced liver injury. In contrast, the cholestatic cluster, defined by elevated ALP, may indicate biliary obstruction or other cholestatic disorders. The normal cluster represents individuals with minimal or no biochemical abnormalities. This stratification highlights the potential of cluster analysis as a tool for personalized medicine, enabling clinicians to tailor diagnostic and therapeutic approaches based on individual biochemical profiles[25].

The integration of multiple statistical techniques in this study represents a significant advancement over traditional approaches that rely solely on descriptive or univariate analyses. By combining ANOVA, regression modeling, data transformation, outlier diagnostics, and clustering, the study provides a comprehensive framework for understanding liver enzyme dynamics. This multidimensional approach enhances the depth and reliability of the findings, making them more applicable to real-world clinical scenarios.

From a broader perspective, the findings of this study contribute to the growing body of evidence supporting the use of advanced analytics in clinical biochemistry. As healthcare increasingly moves toward data-driven decision-making, the ability to analyze complex datasets and extract meaningful insights becomes increasingly important[26]. The application of these techniques to liver enzyme analysis represents a step toward more sophisticated and precise diagnostic tools.

Despite its strengths, the study has several limitations that should be acknowledged. The relatively small sample size may limit the generalizability of the findings, and the cross-sectional design precludes the assessment of temporal changes in enzyme levels. Additionally, the absence of detailed clinical data, such as underlying diagnoses or medication use, limits the ability to correlate biochemical findings with specific disease states[27]. Future studies should aim to address these limitations by incorporating larger, more diverse populations and longitudinal designs.

Another important consideration is the potential influence of external factors such as diet, physical activity, and environmental exposures on liver enzyme levels. These variables were not accounted for in the present study but may contribute to the observed variability. Incorporating such factors into future analyses could provide a more comprehensive understanding of liver enzyme regulation[28].

In conclusion, this study demonstrates that liver enzyme variability is influenced by a combination of demographic and biochemical factors, with ALP emerging as the most variable and clinically informative

parameter. The identification of age and AST as significant predictors of ALP highlights the interconnected nature of hepatic processes, while cluster analysis reveals distinct phenotypic patterns that may have important clinical implications. The use of advanced statistical techniques enhances the robustness and interpretability of the findings, supporting their relevance in both research and clinical settings.

CONCLUSION

The present study provides a comprehensive evaluation of liver enzyme dynamics using both conventional and advanced statistical approaches, offering valuable insights into the variability and interrelationships of AST, ALT, and ALP. Among the parameters analyzed, alkaline phosphatase (ALP) emerged as the most variable enzyme, demonstrating significant differences across age groups and highlighting its sensitivity to physiological and possibly subclinical variations.

The findings indicate that gender does not significantly influence liver enzyme levels within this cohort, suggesting that biochemical variability is largely independent of sex in relatively homogeneous populations. In contrast, age plays a measurable role, particularly in relation to ALP, which showed a declining trend with increasing age. This underscores the importance of considering age-specific variations when interpreting liver enzyme values in clinical practice.

Multivariate regression analysis further identified age and AST as independent predictors of ALP, emphasizing the interplay between hepatocellular and cholestatic processes. The positive association between AST and ALP suggests that liver enzyme abnormalities may not occur in isolation but rather reflect overlapping pathophysiological mechanisms. This reinforces the need for integrated interpretation of liver function tests rather than reliance on individual markers.

The application of log transformation effectively addressed the skewed distribution of ALP, improving the validity and reliability of statistical analyses. Additionally, outlier diagnostics confirmed the robustness of the findings, indicating that the observed associations were not disproportionately influenced by extreme values.

Importantly, cluster analysis revealed three distinct biochemical phenotypes—normal, hepatocellular, and cholestatic patterns—highlighting the heterogeneity of liver enzyme profiles within the study population. This phenotypic classification provides a more nuanced understanding of biochemical variability and may have significant implications for clinical stratification and personalized diagnostic approaches.

Overall, this study demonstrates that the integration of advanced statistical techniques enhances the interpretability and clinical relevance of liver enzyme analysis. The findings support the concept that liver enzyme variability is multifactorial and best understood through a multidimensional analytical framework. Future research with larger sample sizes and incorporation of clinical correlates is warranted to further validate these observations and expand their applicability in clinical practice.

Declarations

Ethics Approval Statement: The study was conducted in accordance with the ethical standards of the institutional research committee. Ethical approval was obtained from the Institutional Ethics Committee (IEC) of AIPH University, Bhubaneswar, Odisha, India. Participant Consent Statement: Informed consent was obtained from all individual participants included in the study. In cases where applicable, the requirement for additional consent for publication was waived by the Institutional Ethics Committee.

Funding: This research was **self-funded**, and no external financial support was received.

Authors and Affiliations

1. **Professor & Director**, AIPH University, Jatni, Khorda, Odisha, India
   Email: directoriqac@aiph.ac.in
2. **Professor, Department of Physiotherapy**, AIPH University, Jatni, Khorda, Odisha, India
   Email: jtiwari@aiph.ac.in

Corresponding Author

Correspondence to: **Vikas Tiwari** (Email: vikas.tiwari.vats@gmail.com)

Conflict of Interest

The authors declare **no conflict of interest**.

References

1. Pratt DS, Kaplan MM. Evaluation of abnormal liver-enzyme results in asymptomatic patients. *N Engl J Med*. 2000;342(17):1266–1271.
2. Giannini EG, Testa R, Savarino V. Liver enzyme alteration: a guide for clinicians. *CMAJ*. 2005;172(3):367–379.
3. Elinav E, Ackerman Z, Maaravi Y, Ben-Dov IZ, Ein-Mor E, Stessman J. Low alanine aminotransferase activity in older people is associated with greater long-term mortality. *J Am Geriatr Soc*. 2006;54(11):1719–1724.
4. Kim HC, Nam CM, Jee SH, Han KH, Oh DK, Suh I. Normal serum aminotransferase concentration and risk of mortality from liver diseases: prospective cohort study. *BMJ*. 2004;328(7446):983.
5. Piton A, Poynard T, Imbert-Bismut F, Khalil L, Delattre J, Pelissier E, et al. Factors associated with serum alanine aminotransferase activity in healthy subjects: consequences for the definition of normal values. *Hepatology*. 1998;27(5):1213–1219.
6. Kwo PY, Cohen SM, Lim JK. ACG clinical guideline: evaluation of abnormal liver chemistries. *Am J Gastroenterol*. 2017;112(1):18–35.

7. Sattar N, Scherbakova O, Ford I, O'Reilly DSJ, Stanley A, Forrest E, et al. Elevated alanine aminotransferase predicts new-onset type 2 diabetes independently of classical risk factors. *Diabetes*. 2004;53(11):2855–2860.
8. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med*. 2004;66(3):411–421.
9. Osborne JW. Improving your data transformations: applying the Box-Cox transformation. *Pract Assess Res Eval*. 2010;15(12):1–9.
10. Cook RD. Detection of influential observation in linear regression. *Technometrics*. 1977;19(1):15–18.
11. Everitt BS, Landau S, Leese M, Stahl D. *Cluster Analysis*. 5th ed. Chichester: Wiley; 2011.
12. Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS, Seeff LB. Diagnosis and monitoring of hepatic injury: I. Performance characteristics of laboratory tests. *Clin Chem*. 2000;46(12):2027–2049.
13. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372(9):793–795.
14. Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall; 1991.
15. Ruhl CE, Everhart JE. Determinants of the association of overweight with elevated serum alanine aminotransferase activity in the United States. *Gastroenterology*. 2003;124(1):71–79.
16. Piton A, Poynard T, Imbert-Bismut F. Gender differences in liver enzyme levels: a population-based study. *Hepatology*. 1998;27(5):1213–1219.
17. Tonelli M, Curhan G, Pfeffer M, Sacks F, Thadhani R, Melamed ML, et al. Relation between alkaline phosphatase, serum phosphate, and all-cause or cardiovascular mortality. *Circulation*. 2009;120(18):1784–1792.
18. Fraser CG. Biological variation in clinical chemistry: an update. *Ann Clin Biochem*. 2001;38(Pt 3):197–208.
19. Kim HC, Jee SH, Han KH, Oh DK, Suh I. Serum aminotransferase concentration and risk of mortality from cardiovascular disease. *BMJ*. 2004;328:983.
20. Sherlock S, Dooley J. *Diseases of the Liver and Biliary System*. 12th ed. Oxford: Blackwell Publishing; 2011.
21. Friedman LS. Liver, biliary tract, and pancreas disorders. In: Goldman L, Schafer AI, editors. *Goldman-Cecil Medicine*. 25th ed. Philadelphia: Elsevier; 2016. p. 1000–1015.
22. Osborne JW. Best practices in data transformation: improving normality and interpretation. *J Mod Appl Stat Methods*. 2002;1(2):1–12.
23. Barnett V, Lewis T. *Outliers in Statistical Data*. 3rd ed. Chichester: Wiley; 1994.
24. Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognit Lett*. 2010;31(8):651–666.
25. Nicholson JK, Lindon JC, Holmes E. Metabonomics: understanding the metabolic responses of living systems. *Xenobiotica*. 1999;29(11):1181–1189.
26. Collins FS, Hamburg MA. First FDA authorization for next-generation sequencer. *N Engl J Med*. 2013;369(25):2369–2371.

27. Altman DG, Bland JM. Statistics notes: absence of evidence is not evidence of absence. *BMJ*. 1995;311(7003):485.
28. Fraser CG. Inherent biological variation and reference values. *Clin Chem Lab Med*. 2004;42(7):758–764.