RESEARCH ARTICLE

Hyperlipidaemia and its future prospective in district Gujrat, Pakistan: a retrospective study

Muhammad Rashid, Farwa Sultan

Abstract

Objective: To evaluate the prevalence of hyperlipidaemia with respect to age and gender, and to assess future prospects.

Method: The retrospective study was conducted in October 2024 at Dr Abdul Rauf Laboratory, Gujrat, Pakistan, and comprised data from January 1, 2020, to September 30, 2024, of patients who had undergone lipid profile test. To assess prospects for the next 10 years, Seasonal Auto-Regressive Integrated Moving Average model was used by using Jupyter Notebook of Anaconda Navigator 1.9.12 that was installed in Microsoft window 7(32 bit).

Results: Of the 4,004 patients, 2,338(58.4%) were male and 1,666(41.6%) were female. The overall prevalence of hyperlipidaemia was 3286(82.1%). The prevalence of hyperlipidaemia among male patients was high throughout the time as 266(86.9%) in 2020, 384(84.4%) in 2021, 444(86.7%) in 2022, 500(83.8%) in 2023 and 408(87.2%) in 2024. In contrast, female patients exhibited comparatively lower prevalence rates as 146(81.1%) in 2020, 298(74.7%) in 2021, 279(74.8%) in 2022, 308(77.4%) in 2023 and 253(80.1%) in 2024. In males age group 41-50 years highly effected with 564(88.3%) while in females age group 51-60 years with 366(80.6%). The prevalence of high level of triglycerides, cholesterol, and low-density lipoprotein, and low level of high-density lipoprotein was 2364(59%), 1454(36.3%), 1172(29.3%) and 2022(50.5%) respectively. Level of triglycerides and high-density lipoprotein was higher in males 1454(62.2%), 1371(58.6%) compared to females 910(54.6%), 651(39.1%) while the level of cholesterol and low-density lipoprotein was higher in females 618(37.1%), 497(29.8%) compared to males 836(35.8%), 675(28.9%). The Seasonal Auto-Regressive Integrated Moving Average model suggested that hyperlipidaemia over the next 10 years in Gujrat will fluctuate between 80% and 90%.

Conclusion: All stakeholders need to play their role in controlling the spread of life-threatening hyperlipidaemia.

Key Words: Hyperlipidaemia, Triglycerides, Cholesterol, High-density lipoprotein, Low-density lipoprotein, Heart diseases, Strokes.

(JPMA 76: 1069; 2026) **DOI: https://doi.org/10.47391/JPMA.30013**

Introduction

Hyperlipidaemia nowadays is a major public health issue that is becoming more common with increasing prevalence throughout the world. Hyperlipidaemia is diagnosed when one or all lipid or lipoprotein levels increase in blood compared to their respective normal values.[1] In primary hyperlipidaemia, some genetic factors are involved in abnormal lipid profile, while in secondary hyperlipidaemia, external factors, like diabetes, hypertension, chronic renal and liver diseases, hypothyroidism, certain medications, unhealthy diet and poor lifestyle, are involved.[2] Hyperlipidaemia is classified into four different clinical subtypes; hypercholesterolemia (elevated level of total cholesterol [TC]), hypertriglyceridemia (elevated level of triglyceride [TG]), mixed hyperlipidaemia, or combined hyperlipidaemia (hypercholesterolaemia, hypertriglyceridaemia with elevated level of low-density lipoprotein cholesterol [LDL-C] and lower level of high-density lipoprotein cholesterol [HDL-C]).[3,4]

According to the expert panel of the National Cholesterol Education Programme (NCEP) final report in Adult Treatment Panel III (ATP-III), lipid profile is measured in mg/dl terms, and hyperlipidaemia is defined as TC >200, TG >150, LDL-C >130 and HDL-C <40.[5] Abnormal lipid profile can be responsible for atherosclerotic plaques or atherosclerotic cardiovascular disease (ASCVD), which is ultimately responsible for peripheral artery disease (PAD), coronary artery disease (CAD) and brain stroke that are among the leading causes of death.[6,7] Nowadays, atherosclerosis events are abruptly increasing around the globe. In the United States, from 2017 to 2020, high TC levels were observed in around 25 million adults.[8] In Pakistan, the prevalence of dyslipidaemia is 96% which is a very alarming situation, and the percentage of hypercholesterolaemia is high in the Punjab province of

Department of Biosciences, Grand Asian University Sialkot, Sialkot, Pakistan.
Correspondence: Muhammad Rashid.
Email: muhammad.rashid@gaus.edu.pk
ORCID ID: 0000-0002-2612-7048

Submission complete: 29-01-2025 **First Revision received:** 16-05-2025
Acceptance: 25-02-2026 **Last Revision received:** 24-02-2026

Pakistan (41.6%) compared to the province of Balochistan (22.7%).[9] About one-third of all deaths globally are caused by ASCVD, of which myocardial infarction (MI) and stroke account for the majority (85%).[10] Globally 4.40 million people had died from elevated levels of LDL-C till 2019.[11] In Pakistan, a critical gap exists in the granular, location-specific epidemiological data on hyperlipidaemia. To our knowledge, no comprehensive analysis has been conducted at the district or provincial levels to accurately quantify the prevalence of hyperlipidaemia, associated mortality and future disease burden. The current study was planned to fill the gap in literature by determining the prevalence of hyperlipidaemia in a district of Punjab province, and to develop a predictive machine learning model to project future trajectory.

Materials and Methods

The retrospective study was conducted in October 2024 at Dr Abdul Raouf Laboratory, Gujrat, Pakistan, and comprised data from January 1, 2020, to September 30, 2024, of patients who had undergone lipid profile test. Reports without gender and age information were excluded and remaining with complete information included in this analysis. Gujrat is a district of the province Punjab, and it consists of Gujrat, Kharian and Sarai-Alamgir tehsils. According to the 2023 census, the overall population of the district was 3,219,375 with a 2.63% annual growth rate. In 2024, the district population's approximate size was about 3,304,044.[12]

Fasting samples of lipid profile were collected after written permission from the Research and Development (R&D) department of the laboratory, which has a number of hospital-based collection centres within the district. Ethical approval was also taken from institutional review board (GAUS/2024/001). The samples were collected using non-probabilistic consecutive sampling technique. The sample size was calculated using the following formulas with 2% margin of error: [13]

Margin error (e) = $$\mathrm{z} \times \sigma/\sqrt{n}$$

Sample size = $$\frac{\frac{z^2 X p(1-p)}{e^2}}{1+(\frac{z^2 X p(1-p)}{e^2 N})}$$

Within the formula, z was the z-score (99% Confidence level (C.L) z-score = 2.58), sigma (σ) was the standard deviation of population, n was the number of samples, p was population proportion (0.5), e was the margin of error, and N was the population size.

Data was stored in Microsoft Excel 2016. The future prospect of hyperlipidaemia prevalence in the district was explored using a machine learning model technique. Seasonal Auto-Regressive Integrated Moving Average (SARIMA) was used for prediction analysis (time series dataset).[14] The coding of the model was created by using Jupyter Notebook of Anaconda Navigator 1.9.12. To run the model, its graphical user interface (GUI) was installed in Microsoft Windows 7 (32-bit).

Important library codings (Supplementary file 1) were imported, and data cleaning and processing were done to check the duplicate and missing values.[14] After data visualisation, additive decomposition method (Supplementary file 1) was used to check the seasonality and trend. Augmented Dickey Fuller (ADF) tests and Rolling Mean and Rolling Standard Deviation methods were used to check out the stationary of the dataset (Supplementary file 1). In ADF, the null hypothesis (H0) was that p>0.05 would suggest the dataset was non-stationary. The dataset was stationary (p<0.05).

SARIMA model selection was done by auto.arima method, which generated automatically the best and fitted models.[14] From these auto-generated models, the model that had the minimum value of Akaike's Information Criterion (AIC) was selected.[14] AIC is an estimator of prediction error, with low AIC value indicating a good model. The selected model was only to be acceptable if its residues behaved like white noise. The white noise is defined as a sequence of residues that should be mean of residuals about zero, normally distributed, constant variance and uncorrelated (independent or no autocorrelation). To establish acceptability, two different approaches were used; the first from the diagnostic test (graphical visualisation), and the second from the Ljung-Box and Jarque-Bera (JB) tests to find the value of Prob (Q) and Prob (JB). The residual test statistics were analysed in the result summary.[15] In the Ljung-Box test, the null hypothesis was that there is no correlation in the residual values, and the null hypothesis would be accepted if Prob (Q) value was >0.05. The alternative hypothesis was that there is a correlation between values if Prob (Q) value was <0.05. In the JB test, the null hypothesis was that residual values are normally distributed if Prob (JB) value was >0.05. Both the assumptions met the p>0.05 condition to qualify for further analysis.

To check the accuracy of the fitted and selected model, "one step ahead" method was used 60% from one side tail and mean absolute percentage error (MAPE) value. The fitted and accurate selected model was used to predict the future prospective of hyperlipidaemia Percentage for the next 10 years in the district.

Results

Of the 4,004 patients, 2,338(58.4%) were male and 1,666(41.6%) were female. The overall prevalence of hyperlipidaemia was 3286(82.1%). In 2020, overall annually hyperlipidaemia prevalence was 412(84.77%), in 2021, 682(79.86%), in 2022, 723(81.69%), in 2023, 808(81.21%) and in 2024, 661(84.31%). The prevalence of hyperlipidaemia among male patients was high as compared to female from 2020 to 2024 as 266(86.9%) in 2020, 384(84.4%) in 2021, 444(86.7%) in 2022, 500(83.8%) in 2023 and 408(87.2%) in 2024. In contrast, female patients exhibited comparatively lower prevalence rates as 146(81.1%) in 2020, 298(74.7%) in 2021, 279(74.8%) in 2022, 308(77.4%) in 2023 and 253(80.1%) in 2024 (Supplementary file Figure 1). Age and gender were strongly associated with hyperlipidaemia. Male age group 41-50 years and female age group 51-60 years had the highest percentage of hyperlipidaemia with 564(88.3%) and 366(80.6%) respectively (Supplementary file Table 1). The overall prevalence of lipid profile parameters TG, TC, LDL-C and HDL-C was 2364(59%), 1454(36.3%), 1172(29.3%) and 2022(50.5%) respectively (Supplementary file Figure 2). Hypertriglyceridaemia in males compared to females was noted in 204(66.7%) and 120(66.7%) subjects in 2020, 311(68.4%) and 222(55.6%) in 2021, 317(61.9%) and 201(53.9%) in 2022, 361(60.5%) and 213(53.5%) in 2023, and 261(55.8%) and 154(48.7%) in 2024. Higher levels of TC in males were noted in 114(37.3%) subjects in 2020, 210(46.2%) in 2021, 187(36.5%) in 2022, 171(28.6%) in 2023 and 154(32.9%) in 2024. The corresponding values for females were 74(41.1%), 183(45.9%), 132(35.4%), 118(29.7%) and 111(35.1%), respectively (Figure 1). The overall prevalence of hypertriglyceridaemia and TC were 324(66.7%) and 188(38.7%) in 2020, 533(62.4%) and 393(46.0%) in 2021, 518(58.5%) and 319(36.0%) in 2022, 574(57.7%) and 289(29.1%) in 2023, and 415(52.9%) and 265(33.8%) in 2024.

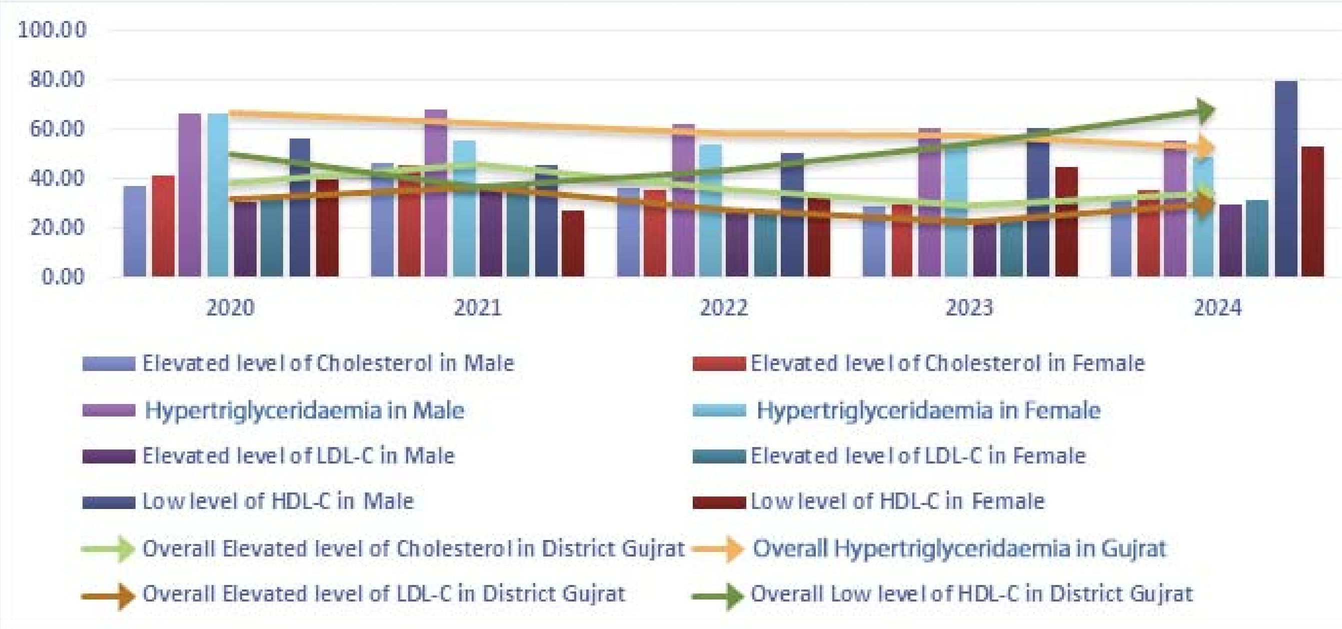


Figure-1: Hyperlipidaemia percentage of both genders along with overall annual percentage.

Supplementary Table-1: High percentages of hyperlipidemia in different age groups of males and females

Years	The annual pattern of Hyperlipidaemia vs Age groups			
	Male		Female	
2020	31-40 n-60 (20.48%)	41-50 n-73(24.91%)	41-50 n-28 (10.98%)	51-60 n-50 (19.61%)
2021	41-50 n-108 (16.95%)	51-60 n-88 (13.81%)	41-50 n-79 (14.91%)	51-60 n-85 (16.04%)
2022	41-50 n-115(15.11%)	51-60 n-117 (15.37%)	41-50 n-66 (13.92%)	51-60 n-75 (15.82%)
2023	31-40 n-134 (15.88%)	41-50 n-143 (16.94%)	41-50 n-75 (13.51%)	51-60 n-87 (15.68%)
2024	31-40 n-114(16.62%)	41-50 n-125 (18.22%)	41-50 n-81 (18.49%)	51-60 n-69(15.75%)

Age Groups with two high percentages of hyperlipidaemia
*n is number of hyperlipidaemia patients.

The overall prevalence of low HDL-C and LDL-C were 244(50.2%) and 154(31.7%) in 2020, 316(37.0%) and 314(36.8%) in 2021, 383(43.3%) and 241(27.2%) in 2022, 540(54.3%) and 225(22.6%) in 2023, and 539(68.8%) and 238(30.4%) in 2024. In males, the prevalence of low HDL-

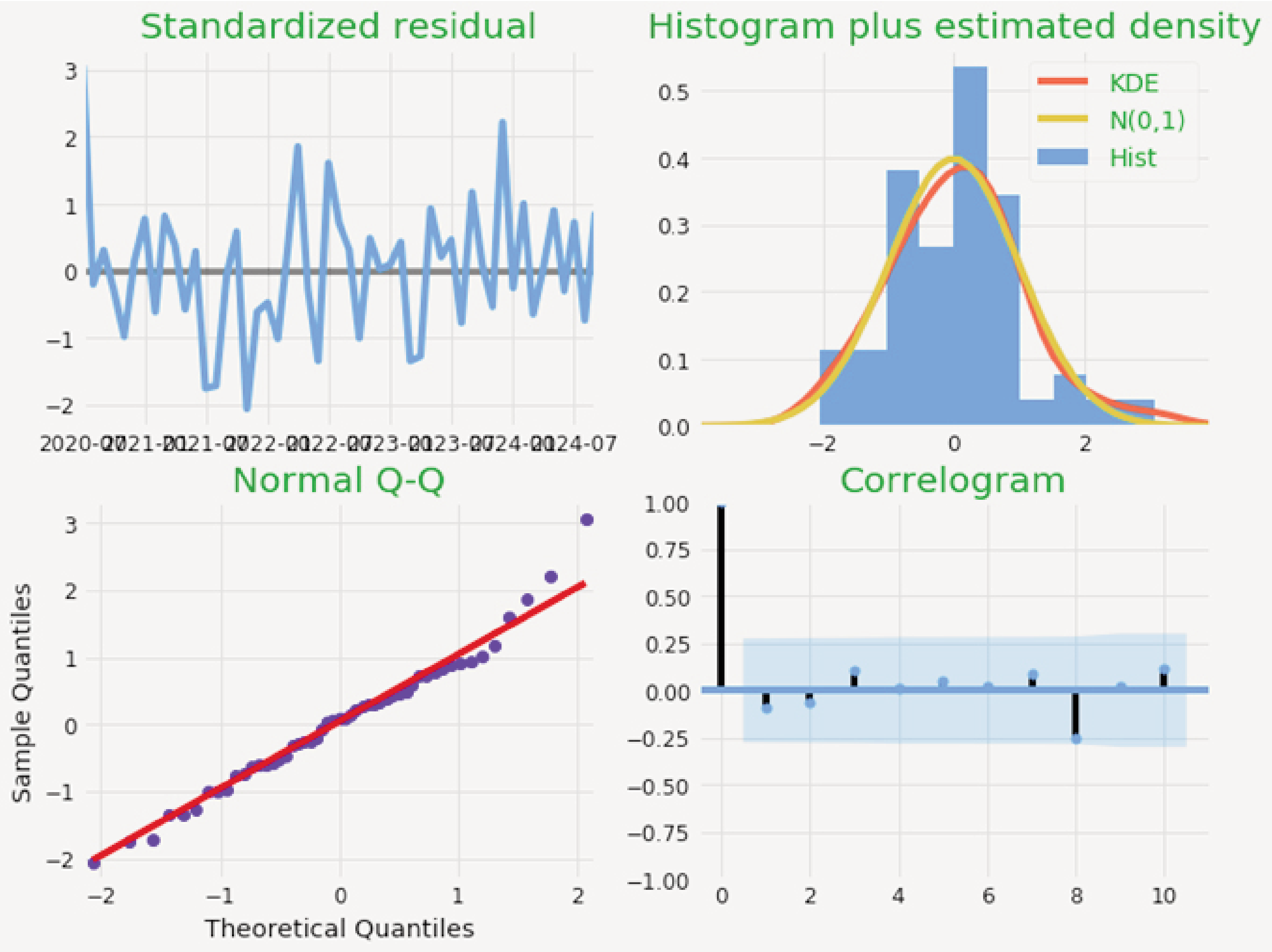


Figure-2: Diagnostic test to check the acceptability of the model used.

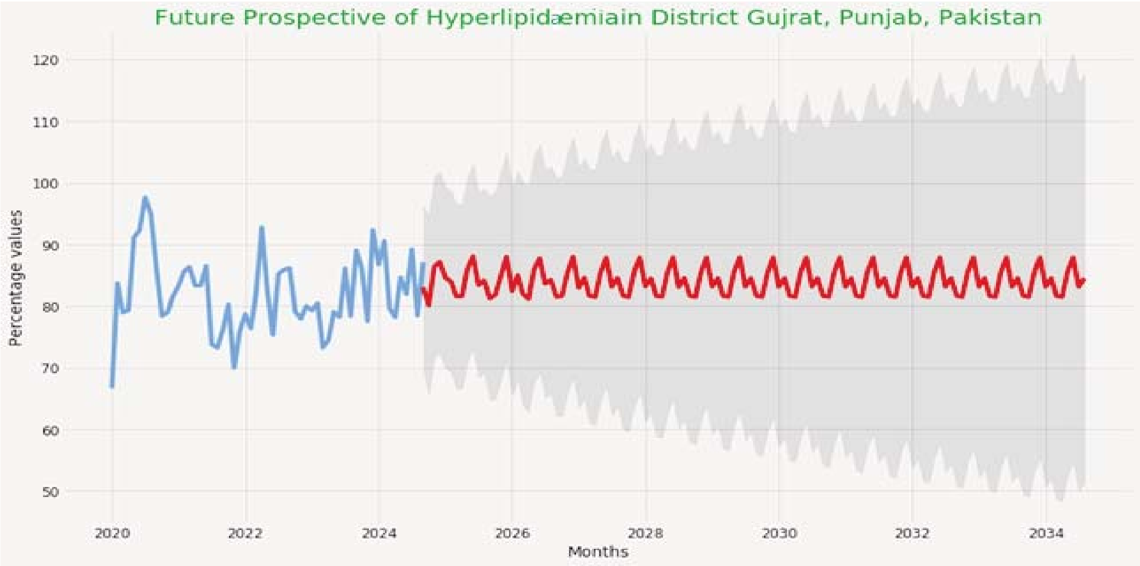
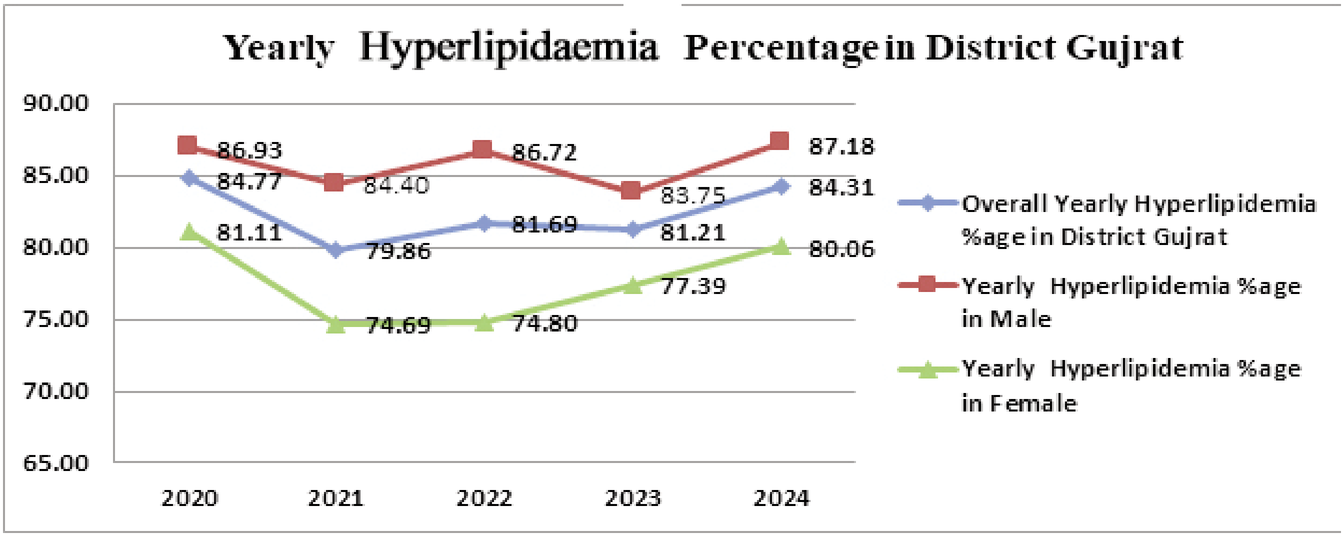
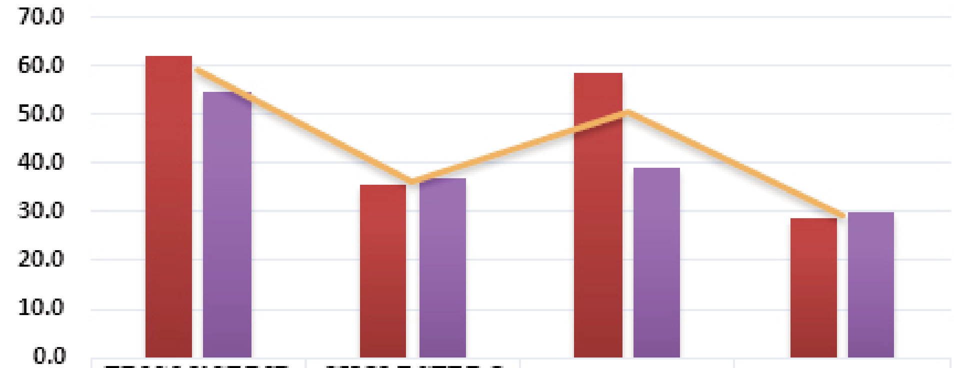


Figure-3: Future prospects (prediction analysis) of hyperlipidaemia in the studied population.



Supplementary Figure-1: From 1st January 2020 to 30th September 2024 Hyperlipidemia percentages of males and females along with overall annually percentage in District Gujrat, Punjab, Pakistan



	TRIGLYCERIDE	CHOLESTEROL	HDL-C	LDL-C
MALE PERCENTAGE	62.2	35.8	58.6	28.9
FEMALE PERCENTAGE	54.6	37.1	39.1	29.8
0VERALL PERCENTAGE	59.0	36.3	50.5	29.3

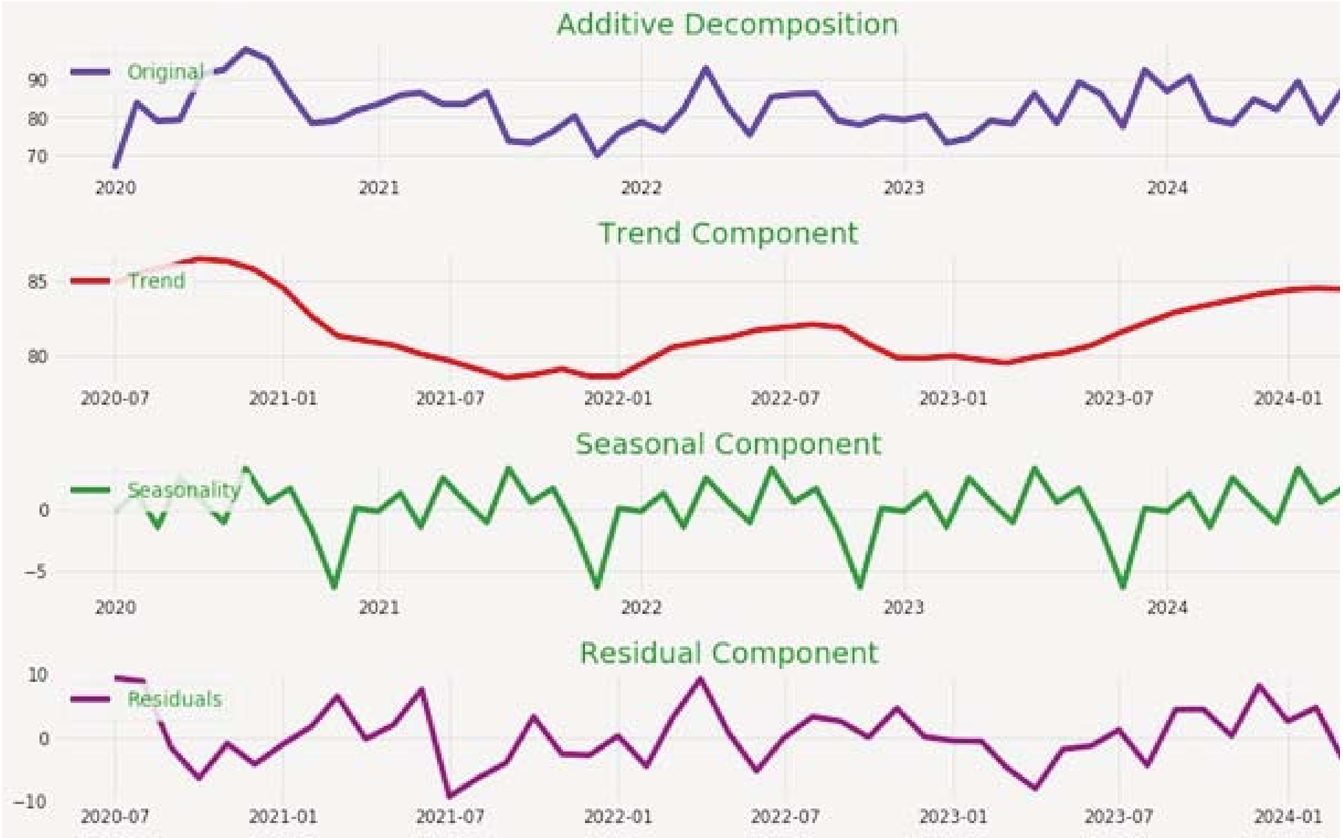
Supplementary Figure-2: Prevalence of lipids profile parameters (TG, Chol, HDL-C, and LDL-C) in District Gujrat.

C levels was consistently higher than in females; 172(56.2%) and 72(40.0%) in 2020, 207(45.5%) and 109(27.3%) in 2021, 257(50.2%) and 126(33.8%) in 2022, 363(60.8%) and 177(44.5%) in 2023, and 299(63.9%) and 152(48.1%) in 2024. Conversely, the LDL-C levels were generally higher in females than in males; 60(33.3%) and 94(30.7%) in 2020, 49(37.3%) and 165(36.3%) in 2021, 98(26.3%) and 143(27.9%) in 2022, 91(22.9%) and 13422.5%) in 2023, and 99(31.3%) and 140(29.9%) in 2024 (Figure 1).

In terms of trends and seasonal components present in the dataset, there were continuously decreasing and increasing trends, while seasonal components with tetra time increasing and decreasing (same pattern) were present in all years (Supplementary file Figure 3).

ADF assessment (p=0.000005) led to the rejection of the null hypothesis. The dataset was stationary because its mean and standard deviation were constant throughout the time (Supplementary file Figure 4). After checking all the parameters of the dataset, the best model was selected with minimum AIC 352.225.

The residuals had constant variance and mean about to

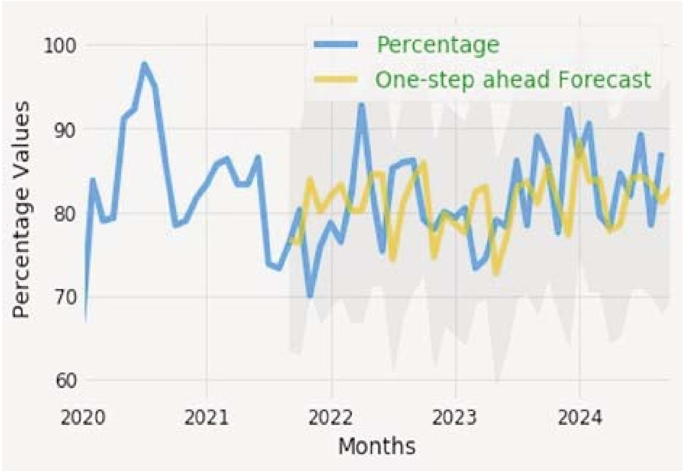


Supplementary Figure-3: Additive decomposition method to check the seasonality and trend of our dataset.

Supplementary Figure-4: Rolling mean & rolling standard deviation methods to check the stationary of the dataset

zero in standardised residues (Figure 2). Histogram and Quantile-Quantile (Q-Q) plot showed that residuals were normally distributed. Correlogram showed no significant correlation in the residual series, indicating that the fitted model residuals satisfied the requirements of the white noise process. In the second approach Prob (Q) value was 0.99, and Prob (JB) value was 0.35. As such, the null hypothesis was accepted in both cases. Before forecast analysis, the accuracy of the best model was checked, and the model followed the same pattern as the original values, indicating that the model was fit for prediction analysis (Supplementary file Figure 5). This was confirmed by MAPE value 6.44 which meant that the model was almost 93.5% accurate.

SARIMAX analysis suggested that hyperlipidaemia over the next 10 years in Gujrat will be fluctuate between 80%



Supplementary Figure-5: One step ahead method to check the accuracy of our fitted and selected model.

and 90% if precautionary measurements did not obtained from relevant departments (Figure 3).

Discussion

The current study showed hyperlipidaemia prevalence of 82.1% in the population of Gujrat, and its percentage in males was higher compared to females. Males and females of age groups 41-50 and 51-60 years, respectively, had a high percentage of hyperlipidaemia, and were at a high risk of CVDs. There is very strong correlation between hyperlipidaemia and CVDs.[16-18] In 2018, a cross-sectional study in Lahore, Pakistan, showed that hyperlipidaemia prevalence in males was high compared to females.[19] This is a global patterns owing to many factors, like smoking, alcohol, high-cholesterol food, testosterone hormones, visceral fat and genetic factors. Besides, the oestrogen hormone in females has a protective effect on lipid profile.[20]

Different countries use different threshold values of TC, TG, HDL-C and LDL-C in relation to ASCVD. A study was conducted in all provinces of Pakistan with data of 10,834 patients, and reported the prevalence of hypercholesterolemia of 39.3%, while the current data from Gujrat district between 2020 and 2024 indicated high level of cholesterol in males and females ranging from 28.6% to 46.2%, which is lower than the overall levels reported from Europe and the United States.[9,21] In the Asia Pacific region, Australia, China, Indonesia, Japan, Malaysia, Singapore, Taiwan and the Philippines have reported the prevalence of abnormal cholesterol range almost equal to that found in the current study, but this region should be a matter of interest because the overall percentage of hypertriglyceridemia is twice as high compared to these Asian countries.[21]

According to the second National Diabetes Survey of Pakistan, hypertriglyceridemia prevalence in Pakistan is 48.9% compared to the 59% noted in the current study.[9,20]

The current data indicates that the annual percentages of low levels of HDL-C are consistently higher in males (58.6%) compared to females (39.1%) in Gujrat. According to 2016-17 dyslipidaemia survey of Pakistan, prevalence of high levels of HDL-C in females was 90% compared to 83.9% in males. In contrast, the prevalence of low HDL-C levels in countries, such as Taiwan, China, Japan, Indonesia and Thailand, is generally lower than in this region, while the Philippines exhibits a notably higher prevalence (71.3%) compared to Gujrat.[21,9] Conversely, the percentage of elevated LDL-C in Gujrat is generally higher among females (29.8%). This trend aligns with observations in Japan and South Korea. Overall, the prevalence of elevated LDL-C is lower in Taiwan, Japan, Indonesia, South Korea, Singapore and China, while in Thailand, it is nearly equal between the genders, while in the Philippines, the rate exceeds that of Gujrat.[21]

The current analysis used the SARIMA model, which was also used to predict seasonal influenza in China.[22] During the coronavirus disease-2019 (COVID-19) pandemic, SARIMA model was used for the prediction of COVID-19 waves, lifecycle and the number of deaths, as well as to estimate the end of the pandemic.[23]

Conclusion

The trend of hyperlipidaemia in Gujrat district was found to be alarming, which calls for due attention.

Disclaimer: None.

Conflict of Interest: None.

Source of Funding: None.

References

1. Dorland WAN. Dorland's Illustrated Medical Dictionary: Dorland's Medical Dictionary Series, 31st ed. Philadelphia, PA: Saunders, 2007; pp 2175.
2. Christie M. Ballantyne, Primary Prevention of Coronary Heart Disease. J Clin Endocrinol Metab 2000;85:2089–92. Doi: 10.1210/jcem.85.6.6642-1.
3. Karr S. Epidemiology and management of hyperlipidemia. Am J Manag Care 2017;23(Suppl 9):s139-48.
4. MedlinePlus. Lipoprotein-a. [Online] 2024 [Cited 2024 October 17]. Available from URL: https://medlineplus.gov/ency/article/007262.htm
5. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002;106:3143-421.
6. Huff T, Boyd B, Jialal I. Physiology, Cholesterol. Treasure Island, FL:

StatPearls Publishing; 2026.
7. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics--2015 update: a report from the American Heart Association. Circulation 2015;131:e29-322. doi: 10.1161/CIR.0000000000000152.
8. Zhao M, Zhang D, Zhang Q, Lin Y, Cao H. Association between composite dietary antioxidant index and hyperlipidemia: a cross-sectional study from NHANES (2005-2020). Sci Rep 2024;14:15935. doi: 10.1038/s41598-024-66922-0.
9. Basit A, Sabir S, Riaz M, Fawwad A; NDSP members. NDSP 05: Prevalence and pattern of dyslipidemia in urban and rural areas of Pakistan; a sub analysis from second National Diabetes Survey of Pakistan (NDSP) 2016-2017. J Diabetes Metab Disord 2020;19:1215-25. doi: 10.1007/s40200-020-00631-z.
10. World Health Organization (WHO). Cardiovascular diseases. [Online] 2024 [Cited 2024 October 23]. Available from URL: https://www.who.int/health-topics/%20cardiovascular-diseases#tab=tab_1.
11. Du H, Shi Q, Song P, Pan XF, Yang X, Chen L, et al. Global Burden Attributable to High Low-Density Lipoprotein-Cholesterol From 1990 to 2019. Front Cardiovasc Med 2022;9:e903126. doi: 10.3389/fcvm.2022.903126.
12. Government of Pakistan, Ministry of Planning Development and Special Initiatives, Pakistan Bureau of Statistics (PBS). 7th Population and Housing Census 2023 (First-ever Digital Census of Pakistan). Islamabad, Pakistan: 2024.
13. SurveyMonkey. Sample size calculator. [Online] 1999-2024 [Cited 2024 September 09]. Available from URL: https://www.surveymonkey.com/mp/sample-size-calculator/.
14. Rashid M, Ismail H. HCV extinction analysis in district Gujrat, Pakistan by using SARIMA and linear regression models. Medicine (Baltimore) 2021;100:e28193. doi: 10.1097/MD.0000000000028193.
15. Dare J, Patrick AO, Oyewola DO. Comparison of stationarity on Ljung box test statistics for forecasting. Earthline J Math Sci 2022;8:325-36. doi: 10.34198/ejms.8222.325336.
16. Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhala N, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet 2010;376:1670-81. doi: 10.1016/S0140-6736(10)61350-5.
17. Sarwar N, Danesh J, Eiriksdottir G, Sigurdsson G, Wareham N, Bingham S, et al. Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies. Circulation 2007;115:450-8. doi: 10.1161/CIRCULATIONAHA.106.637793.
18. Bartlett J, Predazzi IM, Williams SM, Bush WS, Kim Y, Havas S, et al. Is Isolated Low High-Density Lipoprotein Cholesterol a Cardiovascular Disease Risk Factor? New Insights From the Framingham Offspring Study. Circ Cardiovasc Qual Outcomes 2016;9:206-12. doi: 10.1161/CIRCOUTCOMES.115.002436.
19. Zaid M, Hasnain S. Plasma lipid abnormalities in Pakistani population: trends, associated factors, and clinical implications. Braz J Med Biol Res 2018;51:e7239. doi: 10.1590/1414-431X20187239.
20. Pellegrini M, Pallottini V, Marin R, Marino M. Role of the sex hormone estrogen in the prevention of lipid disorder. Curr Med Chem 2014;21:2734-42. doi: 10.2174/0929867321666140303123602.
21. Lin CF, Chang YH, Chien SC, Lin YH, Yeh HY. Epidemiology of dyslipidemia in the Asia Pacific region. Int J Gerontol 2018;12:2-6. Doi: 10.1016/j.ijge.2018.02.010.
22. Cong J, Ren M, Xie S, Wang P. Predicting Seasonal Influenza Based on SARIMA Model, in Mainland China from 2005 to 2018. Int J Environ Res Public Health 2019;16:4760. doi: 10.3390/ijerph16234760.
23. Malki Z, Atlam ES, Ewis A, Dagnew G, Alzighaibi AR, ELmarhomy G, et al. ARIMA models for predicting the end of COVID-19 pandemic and the risk of second rebound. Neural Comput Appl 2021;33:2929-48. doi: 10.1007/s00521-020-05434-0.

AUTHORS' CONTRIBUTIONS:
MR: Concept, design, machine learning analysis, interpretation, editing, writing and final approval.
FS: Data acquisition, interpretation and writing.