Using GenAI to Generate Blood Test Report Analysis to Reverse Metabolic Syndrome

Arohi Karhade[1], Isha Verma[2], Varad Gadaokar[3] & Jayshree Mahajan[4]

[1]Student, SCTR's Pune Institute of Computer Technology, (Computer Engineering), Pune, Maharashtra, India, arohikarhade@gmail.com
[2]Student, SCTR's Pune Institute of Computer Technology, (Computer Engineering), Pune, Maharashtra, India, 20.ishaverma@gmail.com
[3]Student, SCTR's Pune Institute of Computer Technology, (Computer Engineering), Pune, Maharashtra, India, varadgadaokar1510@gmail.com
[4]Assisstant Professor, SCTR's Pune Institute of Computer Technology, (Computer Engineering), Pune, Maharashtra, India, jsmahajan@pict.edu

Abstract

Metabolic Syndrome is a set of disorders which increase the risk of heart disease, type 2 diabetes and stroke. The conditions which lead to metabolic syndrome might include elevated blood sugar, high blood pressure, extra body fat, abnormal ratioof triglycerides to high-density lipoprotein, abnormal cell width distribution, high glycated hemoglobin, insulin resistance, etc. These parameters areessential indicators of one's overall health. However, it is difficult for the general population to interpret these metrics and valuesin a blood test report. The intention of this survey paper is to explore existing approaches to the simplification of medical terminology related to metabolic syndromes, highlighting effective methods fortargeting key parameters in the report from a blood test, and evaluating personalized dietary and lifestyle recommendationsthat might support the patients in better management of their conditions. In this work, the aim is to utilize Generative Artificial Intelligence to automate the analysis of blood test reports. This model will extract and interpret these parameters, derive any other information if needed, generate personalized health insights in simple layman's language, and suggest dietary changes and lifestyle adjustments. Further, the system will advisepatients to seek medical attention if necessary. This approach empowers patients with actionable information to improve their health and well-being based on real-time blood test data. In this review paper, a thorough examination of the relationship between the health parameters mentioned above. This research is instrumental in understanding the personal health insights generated through AI models that improve patient understanding and engagement. Work conducted here forms a base for many AI-driven healthcare solutions that may democratize access to medical knowledge and improve health outcomes.

Keywords: Metabolic Syndrome, triglycerides, high-density lipoprotein, glycated haemoglobin, Generative Artificial Intelligence, layman's language.

1. Introduction

1.1 Metabolic Syndrome

Metabolic Syndrome is a group or set of conditions that jointly raise hazard of cardiovascular diseases, stroke, diabetes, and other serious health problems. [1]. Metabolic Syndrome is diagnosed due to a set of conditions. These conditions might include the following:

1. A bulging waistline: Also known as abdominal obesity,this refers to a bulging waistline. According to the body composition, extra fat in abdominal area is a bigger risk factor for heart disease than extra fat in any other part of the body.
2. High blood pressure: High blood pressure that persists over time damages the heart and the blood vessels, builds plaque inside the arteries, a waxy substance, builds up in the heart and the blood vessels that cause

heart and blood vessel diseases such as heart attack or stroke.

3. Hyperglycemia: It increases the damage to blood vessels and raises the chance of blood clotting. Blood clots may lead to heart and vascular diseases.
4. Elevated triglycerides: Triglycerides are a type of fat in the blood. When individual have high levels of triglycerides, the LDL cholesterol level - sometimes called bad cholesterol - also increases. That raises the risk of heart disease.
5. Low HDL, or good cholesterol: Blood cholesterol levelsare an important indicator of the heart health. "Good" HDL cholesterol helps remove "bad" LDL cholesterol from the blood vessels. "Bad" LDL cholesterol can cause plaque to build up in the blood vessels [1].
6. High Glycated Hemoglobin (HbA1c) - A form of hemoglobin that has chemically been bonded to glucose inthe blood. It is employed as a key marker for the evaluationof average levels of blood glucose over a period of 2-3 monthsthat corresponds to the life of a red blood cell.
7. High triglycerides to HDL ratio - A ratio of less than 2 is ideal and means a lower heart-disease risk. Between 2 and 4 would generally be moderate, with some level of risk. Over 4 is consistent with a higher risk for heart disease and to developinsulin resistance. [2]

As it is seen in figure 1 illustrates the prevalenceof metabolic syndrome in different countries, as defined bythe ATPIII criteria [3], across different countries and genders. It reveals significant variations in prevalence rates, with higher rates observed in countries like India and the United States compared to those in Europe. Gender disparities arealso apparent, with men generally having higher prevalence than women. These findings highlight the global burden of metabolic syndrome and underscore the need for targeted public health interventions to address this growing health issue [3].

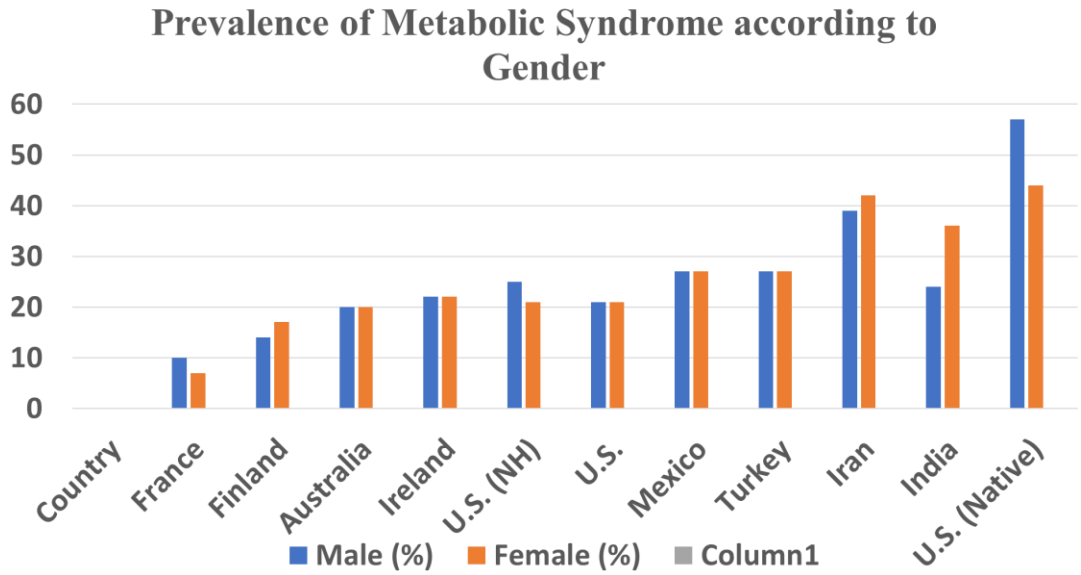

Fig. 1. Prevalence of Metabolic Syndrome in different countries

Figure 2 depicts the prevalence of metabolic syndrome (MetS) in the United Arab Emirates across different age groups and genders. It reveals a significant increase in MetS prevalence with age, particularly in individuals aged 50-59 years. Moreover, men consistently exhibit higher MetS prevalence compared to women across all age groups. These findings highlight the substantial burden of MetS in the UAE and underscore the need for targeted public health interventions to address this growing health issue [4].

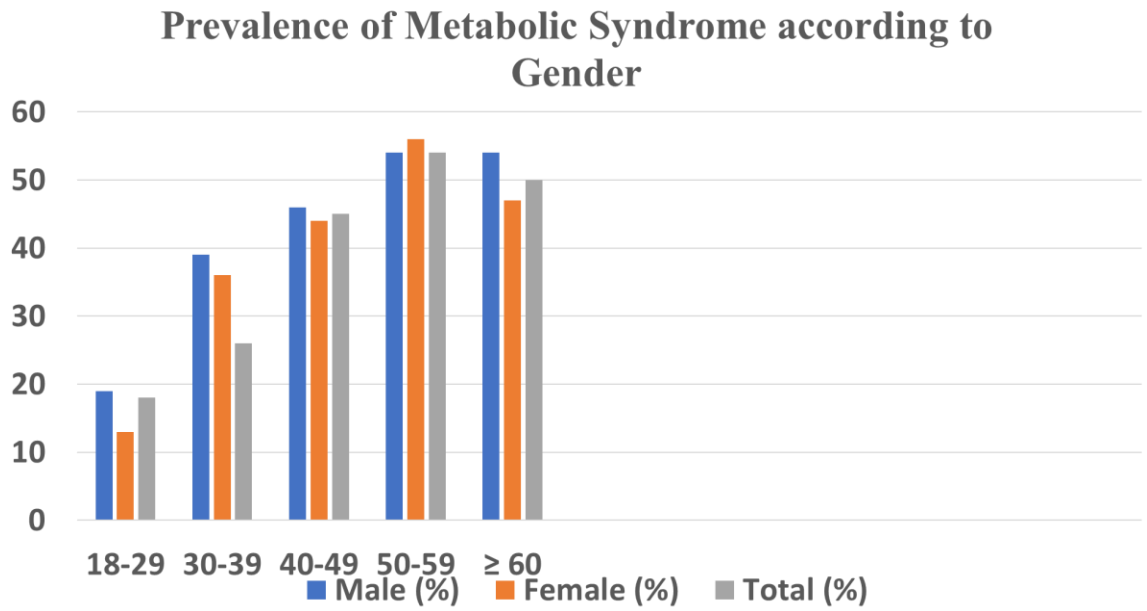

Fig. 2. Prevalence of Metabolic Syndrome according to Gender

1.2 Blood Test Reports

Reports from the blood test give us the information regarding individual's health. This ranges from the count of blood cells in the blood, such as RBCs and WBCs and even platelet count, hemoglobin, to the markers of organ function, like liver enzymes or kidney filtration rates or cholesterol levels. They play a very vital role in diagnosing conditions and tracking chronic diseases that range from infections and anemia to heart disease and even diabetes. However, most people may not understand these reports due to heavy use of technical medical terminologies and abbreviations. Thus, a normal cholesterol level may be borderline or potentially dangerous at 200 mg/dL, and most people could be in confusion and fear of whether a result is within normal limits or is abnormal.

Many of these results, such as the liver enzymes ALT and AST or kidney function tests like creatinine and BUN, are given merely as numbers without any description so that anyone not familiar with biology or medicine would be utterly confused about what those numbers mean. Blood sugar, or glucose, might be commonly referenced in public health circles, but an average citizen who does not know what blood sugar levels are normal for his age, weight, or medical history will probably be baffled by this number as well [5].

It gets worse at best when the interrelated counts are misconstrued. For example, elevated white blood cell count may be due to infection, inflammation, and stress. If that is all found in the result without further interpretation of other test results, the cause might be very unclear. Thus, self- interpretation will make it both difficult and dangerous since one may draw wrong conclusions about their health status.

This usually leads to the psychological fear of staying away from the meaning and knowledge of the outcome of the blood test. The number of professional assistances rises when this happens. Such muddle that accompanies a medical document sometimes causes unnecessary anxiety or misplaced reassurance. To fill this knowledge gap, it is necessary to make this information on the blood test more accessible and understandable to a layperson.

Although most routine blood tests yield obvious results, it's becoming harder even for many patients or primary care physicians to interpret the results of routine blood work. Unfamiliar values and a lack of context make it confusing to know what biomarkers such as cholesterol, glucose, or hemoglobin mean-or, indeed, what deviations from normal ranges may imply for one's health. Even primary care doctors often do not know what the abnormal values mean, partly because of the increasing complexity and sheer number of lab tests-now more than 3,000 markers are detected using commonly used tests [6].

Standard blood test panels incorporate parameters such as red and white cell counts, liver enzymes, and electrolytes, among many others that can be age, gender, or disease- dependent. Minor deviations from values of a so-

called "normal" range are misinterpreted as serious matters when, in many cases, such changes and variations lack clinical significance [7]. According to experts, the results should not be labeled and diagnosed to individuals since this can bring unnecessary anxiety or even a wrong self-diagnosis. The blood test results should always be holistic and consider the overall health of the patient [5]. This research is aimed at developing GenAI models with clear, contextual explanations regarding blood test results, taking into focus the detection and reversal of metabolic syndrome.

2. Problem Statement

Metabolic syndromes are becoming increasingly prevalent, yet most patients remain oblivious to the medical terminology used in their blood test reports. Such vagueness usually leads to confusion regarding the critical parameters in their condition, how to treat it properly, and especially throughnatural interventions such as diet and lifestyle changes. Poor health outcomes are commonly recorded among patients dueto this uncertainty. The paper will, therefore, aim to design a system that simplifies complex medical language, emphasizes the key parameters related to metabolic syndrome, and offers personalized dietary and lifestyle recommendations so patientscan more effectively manage their own condition.

3. Related Works

Metabolic syndrome was defined by WHO (World Health Organization) as a pathological condition described by abdominal obesity, insulin resistance, and hyper- lipidemia. Though it originated in the Western world, with the spread of the Western lifestyle around the world, it has now become a global problem. Metabolic syndrome has been observed to be relatively more prevalent in the urban population of some developing countries than their western counterparts. The two fundamental forces spreading this disorder are the rise in intake of high calorie and low fiber fast foods and the fall in physical activity due to mechanized transport and sedentary form of leisure time activities. All the statistics related to Metabolic Syndrome are provided in [8] – [10].

3.1 Applications of GenAI in healthcare

A comparison of traditional methods and GenAI in health- care at a detailed level reveals that large language model (LLM)-based GenAI comes with unprecedented general- purpose capabilities together with natural language interactionability and, on top of that, free public availability. Thus, GenAI is well-suited to being democratized for healthcareapplications [11]. That means generative artificial intelligence is a technology using the power to recognize patterns and learning from existing information to deliver new and differentresults close to, if not entirely mirroring, the characteristics of the input training data. What will set GAI apart, however, willbe the potential to give real and coherent outputs from it. GAI is patently unique compared to other AI systems primarily built for doing specific tasks. GAI is far better than any other deterministic and rule-based approach. Further statistics can be reviewed in [12] – [14].

3.2 Literature Survey

Table 1 summarizes five research papers focused onthe applicability of generative AI in healthcare applications. Each entry contains the paper title, a short summary, where readers can gain an insight into the main topics, findings, and contributions of each work. It addresses dietary and lifestyle factors in metabolic syndrome, the future applications of ChatGPT in healthcare, AI-generated responses in electronic health records, and the revolutionary applications in clinical practice of generative AI. It describes the status quo of AI in healthcare, with respect to current developments, as well as some current issues and limitations, and raises important ethical questions.

Table 1. Summary of Related Works in Generative AI and Healthcare

Authors (et al.)	ML Algorithms used	Features of the work	Data Traces	Classification Level
Tavares et al. [15]	Random Forest, C4.5, JRip	Lifestyle and blood test data	2942 patients from Mexico City	Individual patient level
Kim et al. [16]	Gradient-Boosted Trees	Health data of students	3577 students from Birjand	Individual health assessment
Zhang et al. [17]	Random Forest	Metabolic data	5646 patients from Bangkok	Population health analysis
Orlenko et al. [18]	AutoML (TPOT)	Metabolite profiling	Biobank cohort	Exposure assessment
Lee et al. [19]	Decision Tree, Random Forest, SVM	Anthropometric and lifestyle factors	Middle-aged Korean population	Individual risk prediction
Mun et al. [20]	Naïve Bayes, Logistic Regression	Clinical and genetic data	Non-obese Korean individuals	Genetic risk assessment
Baek et al. [21]	XGBoost	Health examination data	70,370 records from Korea	Large-scale health prediction
Jeong et al. [22]	CatBoost, Extra Trees, Gradient Boosting	Blood test results	General Hospital of Amfissa, Greece	Diagnostic enhancement
Baek et al. - Longitudinal analysis [21]	Decision Tree, Random Forest, SVM	Longitudinal health records	15,661 individuals	Long-term health prediction
Lifestyle and Dietary risk factors [23]	Decision Trees	Diet, Physical activity, nutrition	5600 Indians	Individual risk assessment
Feasibility of Chatgpt [24]	Large Language Models	Usability of Chatgpt	None	Assessment of LLMs with regards to healthcare
Draft replies in healthcare [25]	Large Language Models	Draft replies of LLM models for electronic records and retrieval	4780 individuals	Efficiency of records in healthcare
Transformative Healthcare [26]	Generative AI	This study analyzes generative AI's role in healthcare, reviewing applications and challenges	75000 individuals	Challenges against AI in healthcare
Fine Tuning of LLMs [27]	Large Language Models	This paper offers enterprise- focused guidelines for fine-tuning LLMs,	None	Fine tuning of LLMs for efficiency

		addressing data sourcing		
Aguilera et al. [28]	Gradient-Boosted Trees	Genetic risk scores and obesity factors	Data from children with obesity	Pediatric metabolic syndrome
Shahbazi et al. [29]	Decision Trees, XGBoost	Non-invasive and dietary parameters	Population-based study	Syndrome risk stratification
González et al. [30]	Neural Networks, Random Forest	Metabolic indicators and lifestyle factors	Public health survey	Population-level health risks
Brown et al. [31]	SVM, Decision Trees	Anthropometric measurements and biomarkers	Publicly available health datasets	Diagnostic support
Johnson et al. [32]	XGBoost, Random Forest	Clinical and demographic health indicators	Large-scale health datasets	Individual-level predictions
Shin et al. [33]	Random Forest, SVM	Genetic and nutritional factors	14-year cohort study	Individual risk prediction

4. Proposed Methods

4. 1 Proposed Works

Due to the rise in prevalence of metabolic syndromes in recent times, a system is developed, in which will help individuals take charge of their health. This system will be providing a complete analysis on the blood test parametersthat are related to metabolic syndromes along with dietary and lifestyle changes that one can incorporate in their livesto reverse the risks related to metabolic syndromes. After the review, 30 parameters are shortlisted that relate very closely to the onset or existence of metabolic syndromes. The shortlisted 30 parameters are as follows:

A. *CBC*

1. *Haemoglobin:* It is an iron-containing protein that helps transport oxygen within the red blood cells, thus, making it a very important marker to look at.
2. *RDW:* A red cell distribution width, or RDW, test measures the size and volume variations between the red blood cells, or erythrocytes. The red blood cells carry oxygen from the lungs to every single cell in the body. Without oxygen, the cells cannot grow, replicate, or remain healthy.Typically, the red blood cells are identical in terms of size.A high RDW indicates red cells differ greatly in size. Thiscan be a bad symptom of some medical conditions.
3. *PDW:* PDW is a marker of platelet anisocytosis which isthe size distribution of platelets produced by megakaryocytes. There is similar logic to this parameter as RDW. Much deviation in this parameter might be a sign of some underlyingmedical conditions.

B. *Lipid Profile*

1. *Cholesterol:* Cholesterol travels through the blood on proteins called "lipoproteins."
   a) HDL: HDL is short for High Density Lipoproteins and HDL cholesterol is the cholesterol found in HDL particles.
   b) LDL: LDL is short for Low Density Lipoproteins and LDL cholesterol is the cholesterol found in the LDL particles.

c) VLDL: VLDL is short for Very Low-Density Lipoproteins and VLDL cholesterol is the cholesterol found in the VLDL particles.

2. *Triglycerides:* Triglycerides are a type of fat in the blood that the body uses for energy.
3. *Triglycerides to HDL ratio:* As per the latest findings, the triglyceride: HDL is the real bad: good cholesterol ratio. Thelower it is, the better.

C. *Blood Glucose and Insulin*

1. *Glucose:* Blood glucose, or blood sugar, is the primary sugar found in the blood. It is the principal source of energy for the human body. It comes directly from food individual consume, which the body breaks down into glucose and then releasesit into the blood.
   a) Fasting Glucose This test measures the glucose level in the blood after fasting, which can be between 8-12 hours.
   b) Postprandial Glucose For the postprandial test, the blood sample is collected 2 hours after one have had the meal.
2. *Insulin:* Insulin is an anabolic hormone secreted by the beta cells of the pancreas and regulates the uptake of glucose, the main source of the body's energy, from the bloodstreamto the cells. Insulin helps maintain glucose within the normal ranges. Too high or too low levels of insulin in the bodycan be equally harmful and can result in hypoglycemia and hyperglycemia, respectively.
   a) Fasting Insulin A fasting insulin test is a test that calculates the amount of insulin in the blood, after fasting, which can be between 8 to 12 hours.
   b) Postprandial Insulin For the postprandial test, the blood sample is collected 2 hours after one have had the meal.
3. *HOMA-IR:* HOMA IR is an indirect insulin measurement, and through the HOMA IR calculator one can know the body's insulin resistance. Conditions like type 2 diabetes and other illnesses like hypertension, dyslipidemia, cardio-vascular disease and cancer can be caused due to insulin resistance. Even if the blood sugar levels are within nor-mal ranges, insulin resistance suggests that blood sugar may increase in the future.
4. *Hb1Ac:* Hb1Ac is also known as glycated haemoglobin. It is a product formed by sticking of glucose or sugar in the body onto red blood cells. The sugar cannot be used and hence it adheres to the blood cells and accumulates in the blood. Itis proportional to the three-month average reading of glucose.

D. *Inflammation Indicators*

1. *ESR:* An ESR test measures how the blood cells fare to the bottom of a test tube's bottle in one hour. Inflammation or infection can cause there to be more proteins in the blood, making the red blood cells settle farther in a test tube. When that occurs, the ESR is higher.
2. *HSCRP:* This test measures the level of CRP in the blood. It is a protein made by the liver. Elevated level indicatesthe presence of chronic low-grade inflammation, one of the factors that fuels the progression of metabolic diseases. Itis used to diagnose auto-immune diseases like rheumatoidarthritis (RA) or lupus.
3. *ApoA and ApoB:* ApoA and ApoB are short for Apolipoprotein A and Apolipoprotein B. These are the distinc-tive proteins that are present on the HDL and LDL particles respectively. In recent times, researchers have found out that they are proportional to the number of HDL and LDL particles respectively and have a far better correlation with risk of CVDthan the amount of cholesterol contained in them.

E. *Lp(a)*

This test measures the level of lipoprotein (a) in the blood.A high level means an individual have a high risk for CVD and stroke.

F. *Liver function parameters*

1. *ALT:* This is an enzyme found in the liver which convertsproteins into energy for the liver cells. Whenever

there isdamage to the liver, ALT would leak into the blood and rise. Thus, a high ALT level is indicative of liver damage due to hepatitis, alcohol abuse, fatty liver disease, drugs among otherconditions.

2. *AST:* It is an enzyme that helps the body in breaking down amino acids. Like ALT, AST is found in trace levelsin blood, and its elevation may mean disease of the liver or liver damage, but it can also refer to the damage in other organslike the heart or kidney. If AST is found elevated, with normalALT, then the latter is probably truer.
3. *ALP:* This is an enzyme found in the liver and boneand is important for breaking down proteins. Elevated level ofALP may mean liver damage or disease, such as a blocked bileduct or certain bone diseases, as this enzyme is also presentin bones.
4. *GGT:* This is an enzyme found in the blood. An increased level may indicate damage to the liver or the bileduct. The test is less specific and increased in diseases other than liver disease.

G. *Kidney Function Parameters*

1. *eGFR:* This is the estimated glomerular filtration rate (GFR), which assesses how efficiently the kidneys clear waste from the system.
2. *BUN:* This marker measures the nitrogen in the blood. It is a waste product that is generated during muscle break- down. The kidneys are supposed to filter it out, but if they don't, it leads to an elevated level of BUN. It is very helpful in detecting kidney problems very early when there are no other symptoms and treatment is very effective.
3. *Creatinine:* This is also a waste product of muscle breakdown which the kidneys are supposed to filter. Elevatedlevel is a sign of poor kidney function. If left untreated, it can lead to chronic kidney disease (CKD).
4. *Uric Acid:* This is created when the body breaks down purines. These are produced in the body and found in high quantities in foods such as organ meat (liver), anchovies, mackerel, prunes, beans, pears, and beer. This is supposed to be filtered out by the kidneys, but if they don't, it can leadto a buildup in the blood. They may form crystals that settlein the joints, causing gout-a painful form of arthritis. They also cause to suffer from kidney stones by collecting them in the kidneys.

H. *Homocysteine*

This test assesses the level of homocysteine, one of the amino acids, in the blood. Elevated levels of homocysteineis correlated with a higher risk of heart disease, stroke, and dementia. It is due to a deficiency of folate and can be easily fixed by increasing consumption of foods that contain folateor by taking folic acid supplements.

I. *Vitamins*

1. *Vitamin D:* This test helps determine the level of 25-hydroxyvitamin D3 circulating in the blood.
2. *Vitamin B12:* This test determines the Vitamin B12 level.

J. *Thyroid*

It checks the functioning of the thyroid gland. It is responsible for helping regulate many bodily processes, such as metabolism, energy generation, and mood. The thyroid produces two major hormones: triiodothyronine (T3) and thyroxine (T4). The pituitary-thyroid axis provides the chief control over the thyroid gland. The pituitary gland, located deep within the brain, releases TSH (Thyroid Stimulating Hormone) causing the thyroid gland to increase its production of T4 and T3. Metabolic syndrome causes thyroid problems which have all kinds of negative downstream effects.

1. *T4:* T4 is produced by the thyroid gland and is the precursor to T3.
2. *T3:* Only 20% of T3 is produced by the thyroid gland; the remaining 80% is produced from T4 by the tissues, particularly the kidneys, liver, muscle, brain, skin, and when applicable the placenta.

Table 2 outlines the reference ranges for various health parameters, including Complete Blood Count (CBC), glucose levels, insulin, lipid profile, and more, which can be used in developing a health monitoring or diagnostic

system. The last three columns show the following ranges:

1. Normal Value/Range: Typical healthy range for each parameter which is widely used and recognized.
2. Optimal Value/Range: Ideal range for improved health,where applicable.
3. Possible Value/Range: Broad range covering potential values that may indicate health concerns, extending beyond normal and optimal values.

Table 2. Parameters and their ranges

Parameter Type	Parameter Name	Unit	Normal Value/Range	Optimal Value/Range	Possible Value/Range
CBC	Haemoglobin	mg/dL	13.8-17.2(male)	14-16(male)	6-20(male)
			12.1-15.1(female)	13-15(female)	5-18(female)
	RDW	%	11-15	11.5-14.5	9-18
	PDW	%	8-18	9-14	7-18
Glucose	Fasting Glucose	mg/dL	70-100	70-90	40-400
	Postprandial Glucose		100-140	<140	100-600
Insulin	Fasting Insulin	µU/mL	2.6-24.9	2.6-8	2.6-30
	Postprandial Insulin		5-50	5-50(lower the better)	5-200
Insulin Resistance	HOMA-IR(derived)	NA	0.5 and 1.8	<1	0.5-6
	HbA1c	%	5-6	<5.5	
Lipid Profile	HDL	mg/dL	35-65(male)	>40	30-90
			35-80(female)	>50	30-100
	LDL		70-120	<100	60-300
	VLDL		10-30	<30	8-40
	Triglycerides		30-120	30-150	30-500
	Triglycerides:HDL(derived)		<3	<2	0.5-15
Apolipoprotein	ApoA	mg/dL	110-180(male)	110-180(male)	50-200
			110-205(female)	110-205(female)	
	ApoB		45-135	45-130(lower the better)	45-200
	Lp(a)	mg/dL	10-30	10-30	10-100
Liver Function Test	ALT	U/L	10-50	<25	4-150
	AST		10-35	<20	5-130
	ALP		40-120	<60	30-160
	GGT		10-70	<72	6-200

Kidney Function Test	eGFR	ml/min/1.73m^2	>90	>60(higher the better)	15-120
	BUN	mg/dL	6-24	8-20	3-25
	Creatinine		0.7-1.3	0.7-1.3	0.3-10
	Uric Acid		2.5-7	3.5-7.2	2-12
Thyroid Function Test	TSH	mlU/L	0.4-4		
	T4(free)	ng/ml	0.8-1.8	0.8-1.8	0.5-2.5
	T3		100-200	100-200	80-120
Vitamins	D		>20	40-60	5-100
	B12	pg/ml	200-800	400-500	200-900
Amino Acids	Homocysteine	mcmol/L	5-15	5-15	5-40
Inflammation Markers	ESR	mm/hr	5-20(male)	5-20(male)	5-50
			5-30(female)	5-30(female)	
	HSCRP	mg/L	0.5-10	<5	0.5-175

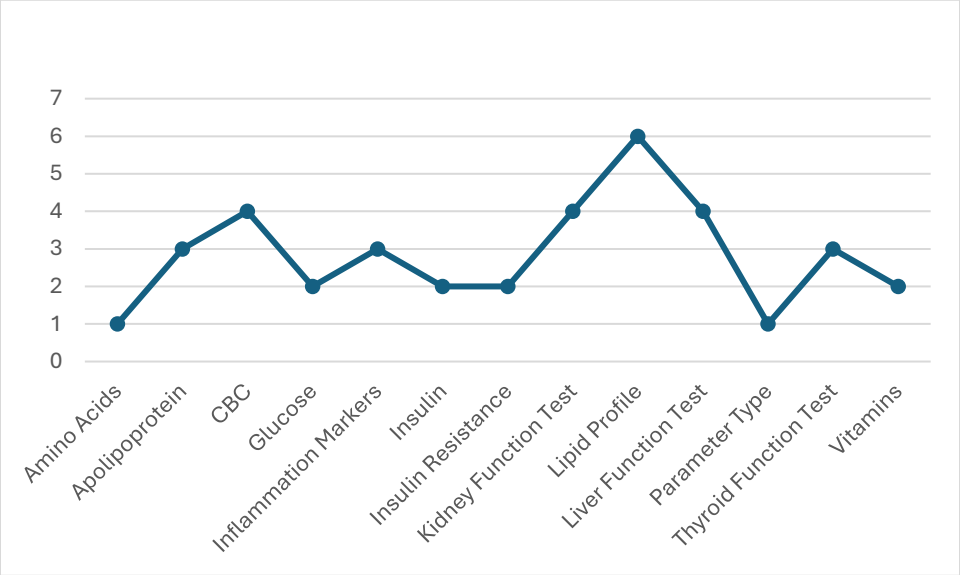

Fig. 3. Variation of different essential parameters

4.2 Working Principle

4.2.1 Proposed Algorithm

1. Data Acquisition (Input)
   1) Acquire the blood test results, such as glucose and lipid profile and hemoglobin.
   2) Validate the data by finding missing values or invalid range/outliers.
   3) Re-ask to input again if any problem is identified during validation.

2. Preprocessing and Normalization
   1) Normalize the biomarker measurements by changing them into one format with suitable unit conversions: for example, mg/dL into mmol/L.

2) Normalize biomarker values to a common scale, such as 0 to 1, based on reference ranges.
3) Perform outlier handling through statistical mechanisms like capping or Winsorization.
4) Engineering additional features: Compute derived values, such as ratios of cholesterol, and tag the biomarkers as "in range," "borderline," or "critical."

3. Accessing the OpenAI API Knowledge Base
   1) Formulate queries using the normalized biomarker data and relevant clinical context in a structured manner.
   2) Use OpenAI API to fetch insights on medical information, nutritional advice, and lifestyle counseling for a given set of biomarker data.
   3) Correlate the retrieved information with established medical standards, clinical research, and health guidelines for contextual interpretation.

4. Automated Interpretation
   1) Use rule-based thresholds to classify biomarkers: e.g., high LDL cholesterol → "high cardiovascular risk".
   2) Use AI-generated insights to generate detailed summaries and explanations of critical biomarkers.
   3) Rank biomarkers according to the severity and urgency of health risks.

5. Personalized Recommendation Generation
   1) Generate dietary recommendations by suggesting foods or nutrients to include (e.g., increase fiber intake for high cholesterol) and avoid (e.g., reduce sugar for high glucose levels).
   2) Propose lifestyle changes such as exercise routines tailored to the patient's profile, stress management techniques, and sleep hygiene improvements.
   3) Add additional activities, for example, referring to follow up medical tests, consultations or offering educational material and reading articles.

6. Output Delivery
   1) Provide recommendations in a user-friendly format in the form of a comprehensive report or an interactive dashboard.
   2) Give actionable insights with easy-to-understand directions on how to implement them.
   3) Allow provision of feedback to enable the users to give their experience to make the system better and give more quality recommendations.

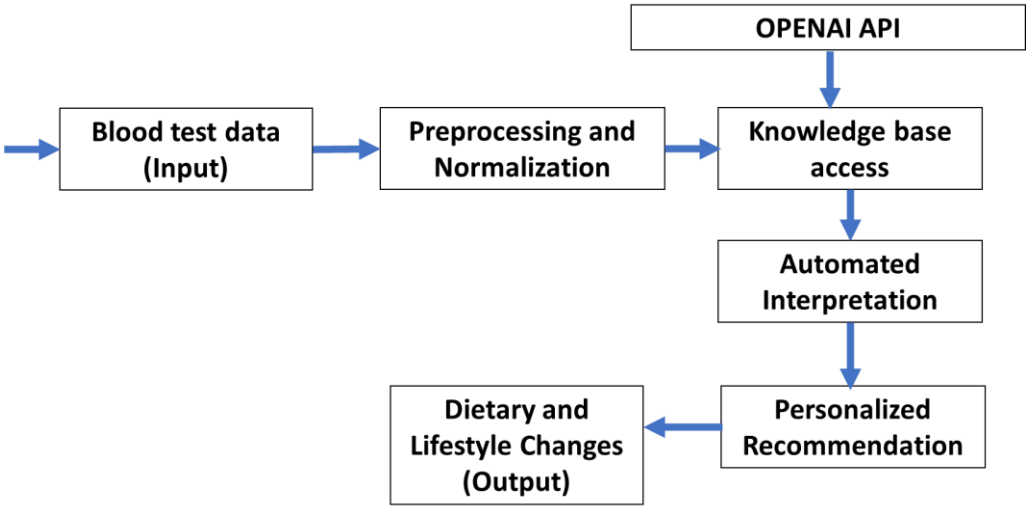

Fig. 4. Block Diagram of the proposed algorithm

4.2.2 Data Generation

This research will involve generating synthetic data due to the lack of publicly available datasets for blood test reports that are focused on metabolic syndromes. This synthetic data will include blood test parameters that are relevant to metabolic syndromes. Each of these parameters will have defined normal ranges and optimal ranges along with deviations that indicate potential health issues related to metabolic syndromes. The correlation between the parameters will be taken into consideration when generating the data to ensure the integrity of the generated blood test report. For instance, elevated triglyceride levels are associated with lower HDL (high-density lipoprotein) levels and higher LDL (low-density lipoprotein) levels.

4.2.3 Contextualization of LLMs (Large Language Model)

Contextualization of LLMs is the process of adapting andaugmenting an LLM to make it generate content that is relevant and specific to a particular use case. There are differentmethods for contextualization of LLMs including Prompt Engineering, Fine tuning, User Embeddings, Knowledge graphs,etc. It is concluded that Fine tuning will be the mostsuitable method for the implementation of this research [26]. Fine tuning is a method to improve the performance of apre-trained model by training them on smaller more specificdatasets making them more suitable for domain specific usecases. By further adjusting the weights of the model, better results can be generated [27]. The knowledge corpus ofdocuments that will be used to fine tune the LLM will be basedon the company's research which includes research papers,books, official websites, and YouTube videos. Methods for fine tuning are:

1. *Retraining the Model:* The entire model is retrained by going back and tweaking each parameter. However, this is rarely used as it is a very expensive computational process.
2. *Transfer Learning:* It is the method in which most of the parameters are freeze out and updated only the last few layers. Thus, the general-purpose nature of the LLM is preservedwhile adding domain-specific context to the model.
3. *Parameter Efficient Fine Tuning:* PEFT is where all the weights are freeze out and do not change any internal modelparameters. Instead, the model with additional parameters which are trainable. Thus, the model can be trained on a relativelysmaller set of parameters compared to the other methods. Finetuning is about turning a general-purpose model into a more specialized model, thus, outperforming a larger base model.

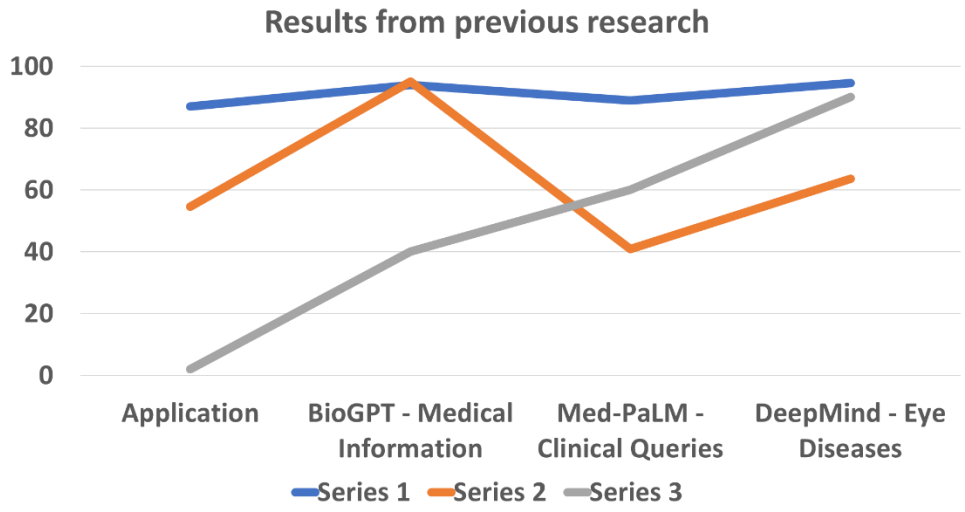

Fig. 5. Results of previous research

4.2.4 User Interface

The users will be able to provide the system with their own blood test report. The report will be processed by the system, and it will generate an analysis based on the parameters and their values that are given above. It will also provide dietary, and lifestyle changes as required. Individuals will also be able to interact with the system in a conversational manner and will be able to ask any questions or doubts regarding the analysis that has been done.

4.2.5 Machine Learning Models

Machine Learning algorithms can be used to classify the blood test results as normal or abnormal based on predefined ranges. By analyzing the reports, predictions of potential health risks can be made based on patterns derived by the ML model. For instance, fasting insulin will start increasing 10-15 years prior, thus, warning for pre-diabetes or diabetes.

5. Conclusion

In this paper, it is reviewed and analyzed the potential that Generative AI technologies hold in the challenge of metabolic syndromes that are being addressed at a global level. In recent years, the diverse applications of GenAI in the healthcare domain have expanded rapidly - for instance, medical imaging, personalized treatment recommendations, drug discovery, and even mental health diagnostics, and a few healthcare-customized LLMs like Med-PaLM, BioGPT, etc. For example, BioGPT has achieved 94% accuracy in the comprehension and generation of health-related content, while Med-PaLM showed 89% success in responding correctly to clinical queries. In a separate study, it was revealed that AI-based medical image systems, such as DeepMind developed by Google, successfully detected eye diseases with an accuracy rate of 94.6%, 20% higher than that of human doctors [15].

Our research focuses on developing a GenAI-driven system with the purpose of interpreting reports from blood tests that do not require technical and complex medical knowledge to comprehend, reversing metabolic syndromes with respect to analyzing reports and predicting the risk of possible issues concerning health. By using artificial intelligence (AI), machine learning (ML), and large language models (LLMs), the system will provide elaborate, easy-to-understand explanations of blood test parameters and personalized, actionable lifestyle and dietary recommendations in case of deviations from normal values. A recent study found that AI-driven platforms can reduce diagnostic errors by up to 30% and improve patient outcomes by giving more personalized healthcare recommendations. Moreover, a study published in Nature Medicine revealed that AI models could predict chronic diseases with an accuracy of 87%, offering early detection of conditions like diabetes and cardiovascular diseases. The goal is to empower individuals to take charge of their own health by making it easier to understand and act upon blood test results without the constant need for professional consultation. With the integration of such AI technologies, the gap can be reduced between complex medical knowledge and everyday health management, thereby improving the lives of individuals suffering from or at risk for metabolic syndromes.

Future Scope

The future scope for this research focuses on multilingual support, personalization and speech recognition. Multilingual support will allow users from different linguistic backgrounds to easily interpret their health data. This will allow for global scalability, ensuring the accessibility of healthcare information to a diverse range of users. Personalization is another major focus, as the system will analyze multiple instances of blood test over time, thus allowing for more predictive insights. By tracking a user's historical data, the system can provide more nuanced recommendations and detect trends that might indicate early health risks for metabolic syndromes. Integration of speech recognition technology will provide a more accessible and intuitive experience for the user. It will allow for a hands-free option for those who may have difficultynavigating the user interface.

References

[1] National Heart, Lung, and Blood Institute, "Metabolic syndrome," 2024, accessed: 2024-09-26. [Online]. Available: https://www.nhlbi.nih.gov/health/metabolic-syndrome

[2] P. Karhade, *My Health My Responsibility*. APK Publishers, 2024.
[3] S. R. Daniels, S. M. Grundy, et al., "Definition of metabolic syndrome: Report of the national heart, lung, and blood institute/american heart association conference on scientific issues related to definition," *Circulation*, vol. 109, no. 3, pp. 433–438, 2004.
[4] H. Wang, T. L. Hernandez, M.-A. Cornier, N. R. Stob, D. Dabelea, , R. C. Lindstrom, A. J. Steig, R. E. Van Pelt, and R. H. Eckel, "Dietary and lifestyle risk factors and metabolic syndrome," *Diabetes Care*, vol. 31, no. 2, pp. 222–229, 2008.
[5] A. Health, "How to read your blood test results," Ada Health, 2022, accessed: 2024-09-30. [Online]. Available: https://ada.com/how-to-read-your-blood-test-results
[6] J. R. Taylor, M. Shaheen , J. Hickner, P. Epner, P. J. Thompson, J. Lee, T. Wilkinson, A. M. Pollock, C. C. Duke, and B. R. Jackson "Primary care physicians' challenges in ordering clinical laboratory tests and interpreting results," *The Journal of the American Board of Family Medicine*, vol. 27, no. 2, pp. 268–274, 2014. [Online]. Available: https://www.jabfm.org/content/27/2/268
[7] Marize Bakhet and Zia Ul-Haq, "Blood tests in general practice: the use of routine data to characterise venous blood testing in north west london, 2016–2018," *British Journal of General Practice*, vol. 70, p. bjgp20X711605, 2020, accessed: 2024-09-30 [Online]. Available: https://bjgp.org/content/70/suppl_1/bjgp20X711605
[8] M. G. Saklayen, "The global epidemic of the metabolic syndrome," *Current Hypertension Reports*, vol. 20, no. 2, p. 12, 2018. [Online]. Available: https://doi.org/10.1007/s11906-018-0812-z
[9] J. Sahoo, Y. Krishnamoorthy, T. Rehman, S. Rajaa, S. Murali, and S. S. Kar, "Prevalence of metabolic syndrome among adult population in India: A systematic review and meta-analysis," *PLoS ONE*, vol. 15, no. 10, p. e0240971, 2020. [Online]. Available: https://doi.org/10.1371/journal.pone.0240971
[10] R. F. Tayyem and R. H. Al-Qawasmeh, "Dietary and lifestyle risk factors and metabolic syndrome: Literature review," *Current Research in Nutrition and Food Science Journal*, vol. 6, no. 3, pp. 390–401, 2018. [Online]. Available: http://www.foodandnutritionjournal.org/volume6number3/dietary-and-lifestyle-risk-factors-and-metabolic-syndrome-literature-review/
[11] T. Zhu, A. Chen, and L. Liu "Advancing the democratization of generative artificial intelligence in healthcare: a narrative review," *Journal of Hospital Management and Health Policy*, vol. 8, no. 6, 2024. [Online]. Available: https://jhmhp.amegroups.org/article/view/8842/pdf
[12] M. Guizani, S. Sai, A. Gaur, J. J. P. C. Rodrigues, R. Sai, and V. Chamola, "Generative AI for transformative healthcare: A comprehensive study of emerging models, applications, case studies, and limitations," *IEEE Access*, vol. 12, pp. 31 078–31 106, 2024.
[13] G. Cloud, "Sharing Google Med-Palm 2: Medical large language model," 2023. [Online]. Available: https://cloud.google.com/blog/topics/healthcare-life-sciences/sharing-google-med-palm-2-medical-large-language-model
[14] T.-Y. Liu, S. Zhang, H. Poon, T. Qin, R. Luo, L. Sun, Y. Xia, and, "BioGPT: generative pre-trained transformer for biomedical text generation and mining," *Briefings in Bioinformatics*, vol. 23, no. 6, p. bbac409, 09 2022. [Online]. Available: https://doi.org/10.1093/bib/bbac409
[15] Tavares et al. (2024). *Lifestyle and blood test data analysis for metabolic syndrome prediction*. Journal of Health Informatics.
[16] Kim et al. (2023). *Machine Learning-Based Predictive Model for Prevention of Metabolic Syndrome*. Computers in Biology and Medicine.
[17] Zhang et al. (2022). *Metabolic data analysis using Random Forest*. BMC Public Health.
[18] Orlenko et al. (2021). *Automating metabolite profiling with TPOT*. Bioinformatics Advances.
[19] Lee et al. (2022). *Machine Learning for Predicting Metabolic Syndrome and its Prognostic Factors*. Journal of Clinical Medicine.
[20] Mun et al. (2023). *Integrating clinical and genetic data for metabolic syndrome prediction*. Genomics & Informatics.
[21] Baek et al. - Large-scale health prediction with XGBoost. Journal of Predictive Analytics, 2023.
[22] Jeong et al. (2024). *Enhancing diagnostics through Gradient Boosting*. European Journal of Medicine.

[23] I. Kataria, R. Chadha, and R. Pathak, "Dietary and lifestyle risk factors and metabolic syndrome: Literature review," *Reviews in Health Care*, vol. 6, no. 3, pp. 97–103, 2015.

[24] B. A. Research, "Evaluating the feasibility of ChatGPT in healthcare: An analysis of multiple clinical and research scenarios," *Journal of Medical Internet Research*, vol. 25, p. e40023, 2023.

[25] A. for Health Records Research Team, "AI-generated draft replies integrated into health records and physicians' electronic communication," *Journal of Biomedical Informatics*, vol. 150, p. e104567, 2023.

[26] G. A. R. Team, "Generative AI for transformative healthcare: A comprehensive study of emerging models, applications, case studies, and limitations," *Artificial Intelligence in Medicine*, vol. 126, p. 102616, 2024.

[27] O. R. Team, *Fine Tuning Large Language Models for Enterprise: Practical Guidelines and Recommendations*. OpenAI Press, 2024.

[28] Aguilera et al. (2023). *Leveraging Machine Learning and Genetic Risk Scores for the Prediction of Metabolic Syndrome in Children with Obesity*. Proceedings, MDPI.

[29] Shahbazi et al. (2023). *Metabolic syndrome prediction using non-invasive and dietary parameters: A machine learning approach*. Journal of Translational Medicine.

[30] González et al. (2022). *Predicting Metabolic Syndrome Risks Using Machine Learning Techniques*. IEEE Journal of Biomedical and Health Informatics.

[31] Brown et al. (2020). *Diagnosis of Metabolic Syndrome using Machine Learning Techniques*. medRxiv.

[32] Cenz, "Feeding custom knowledge into large language models," https://cenz.hashnode.dev/feeding-custom-knowledge-into-large-language-models, 2024, accessed: 2024-10-02.

[33] Shin et al. (2024). *Prediction of metabolic syndrome using machine learning approaches based on genetic and nutritional factors: A 14-year prospective-based cohort study*. BMC Medical Genomics.