CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Development of MetSMarker: An Integrated Platform for Curated Biomarkers of Metabolic Syndrome

by

Muhammad Ammar

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Health and Life Sciences
Department of Bioinformatics and Biosciences

2026

CERTIFICATE OF APPROVAL

Development of MetSMarker: An Integrated Platform for Curated Biomarkers of Metabolic Syndrome

by

Muhammad Ammar

(MBS251010)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Amna Abdul Rahim	MY Uni., Islamabad
(b)	Internal Examiner	Dr. Sohail Ahmad Jan	CUST, Islamabad

Dr. Syeda Marriam Bakhtiar
Thesis Supervisor
August, 2026

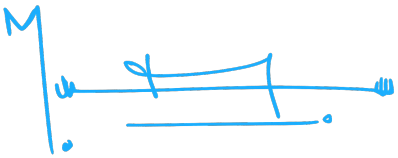
Dr. Syeda Marriam Bakhtiar
Head
Dept. of Bioinfo. and Biosciences
August, 2026

Dr. Sahar Fazal
Dean
Faculty of Health and Life Sciences
August, 2026

Author's Declaration

I, **Muhammad Ammar** hereby state that my MS thesis titled "**Development of MetSMarker: An Integrated Platform for Curated Biomarkers of Metabolic Syndrome**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.



Muhammad Ammar

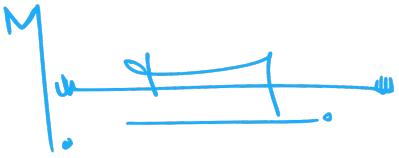
Registration No: MBS251010

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Development of MetSMarker: An Integrated Platform for Curated Biomarkers of Metabolic Syndrome**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.



Muhammad Ammar

Registration No: MBS251010

Acknowledgement

First and foremost, I am deeply grateful to Allah Almighty for blessing me with the strength, patience, wisdom, and guidance to complete this thesis.

I would like to express my sincere gratitude to my supervisor, Dr. Syeda Marriam Bakhtiar, for her invaluable guidance, continuous encouragement, constructive feedback, and unwavering support throughout my MS research. Her mentorship, insightful discussions, and confidence in my abilities have been instrumental in the successful completion of this work.

I also extend my heartfelt appreciation to the members and alumni of the Genetic and Molecular Epidemiology Research Group for fostering a collaborative and inspiring research environment and for their continued contributions to this field over the years. I am especially grateful to Syeda Mohadsa Zainab for her constant support, guidance, and willingness to share her knowledge whenever I needed it. My sincere thanks also go to all the faculty members of the Department of Bioinformatics and Biosciences for their dedication to teaching, mentorship, and for providing a strong academic foundation throughout my studies.

I am especially thankful to Mr. Muhammad Bilal and Mr. Muhammad Qasim for their valuable support during the deployment phase of this research. Their technical assistance, prompt help, and readiness to share their expertise are sincerely appreciated.

Finally, I would like to express my deepest gratitude to my family and friends for their unconditional love, endless encouragement, patience, and unwavering belief in me throughout this journey. Their support has been my greatest source of strength and motivation, and this achievement would not have been possible without them.

(Muhammad Ammar)

Abstract

Metabolic syndrome is a multifactorial disorder characterized by interconnected metabolic abnormalities that increase the risk of cardiovascular diseases, type-2 diabetes, obesity, chronic inflammation, and other chronic disorders. The molecular basis of MetS involves complex genetic, transcriptomic, proteomic, metabolomic, and immunological alterations, emphasizing the need for a comprehensive resource integrating experimentally validated biomarkers. This study aimed to develop MetSMarker, a comprehensive knowledgebase of experimentally validated biomarkers associated with metabolic syndrome. A systematic literature review of PubMed-indexed studies was conducted using predefined inclusion and exclusion criteria. Manual curation of 527 articles identified 207 experimentally validated biomarkers, which were annotated using international databases and standardized nomenclature to ensure consistency, interoperability, and biological relevance. The curated dataset comprised 95 protein, 56 metabolic, 18 clinical, 17 RNA, 14 immunological, and 7 DNA biomarkers, demonstrating the heterogeneity of MetS. Organ-specific annotation identified predominant biomarker involvement in the heart (97), kidney (55), liver (46), pancreas (38), adipose tissue (20), brain (16), and skeletal muscle (14). C-reactive protein and NT-proBNP were the most extensively validated, supported by 13 and 8 independent studies, respectively. MetSMarker enables systematic biomarker exploration through molecular classification, disease- and organ-specific annotation, evidence-based prioritization, comparative analyses, and interactive visualization. These features facilitate molecular interpretation, biomarker research, candidate biomarker prioritization, and systems-level investigation of MetS. MetSMarker provides a standardized, searchable, and analytically rich knowledgebase that integrates heterogeneous biomarker evidence into a single platform. This resource is expected to facilitate biomarker discovery, enhance understanding of MetS pathophysiology, and support translational research toward precision diagnostics and therapeutic development.

Keywords: Metabolic syndrome, MetSMarker, Biomarkers, Knowledgebase, Bioinformatics, Biomarker curation, Precision diagnostics

Contents

List of Figures

List of Tables

Abbreviations

MetS	Metabolic syndrome
WHO	World Health Organization
IDF	International Diabetes Federation
NCEP	National Cholesterol Education Program
ATP III	Adult Treatment Panel III
T2DM	Type 2 diabetes mellitus
BMI	Body mass index
CRP	C-reactive protein
hs-CRP	High-sensitivity C-reactive protein
HbA1c	Glycated hemoglobin
HDL-C	High-density lipoprotein cholesterol
NAFLD	Non-alcoholic fatty liver disease
TNF-α	Tumor necrosis factor-alpha
IL-6	Interleukin-6
IR	Insulin resistance
HOMA-IR	Homeostatic model assessment of insulin resistance
ROS	Reactive oxygen species
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
miRNA	MicroRNA
mRNA	Messenger RNA
SNP	Single nucleotide polymorphism

Chapter 1

Introduction

1.1 Background of Metabolic Syndrome

Metabolic syndrome (MetS) is a complex and multifactorial clinical condition characterized by a cluster of interrelated metabolic abnormalities, including central obesity, insulin resistance, dyslipidemia, and hypertension. These components interact synergistically and contribute to underlying pathophysiological processes such as chronic low-grade inflammation, oxidative stress, and endothelial dysfunction, ultimately increasing the risk of cardiovascular disease and other noncommunicable disorders [1].

The syndrome reflects a convergence of genetic and environmental influences, where lifestyle factors such as poor diet and physical inactivity play a critical role in its development. As a result, MetS has emerged as a significant global public health concern due to its strong association with morbidity and mortality [2]. Metabolic syndrome is of great public health importance because it is prevalent worldwide.

Metabolic syndrome is prevalent in 12.5–31.4% of adults worldwide and is related to risks of cardiovascular disease, diabetes, cancer and mortality in a dose-dependent manner [5]. This broad prevalence within populations and a strong

dose-dependent relationship with mortality highlights the importance of early identification and management of the individual components of the syndrome in preventive and primary care medicine.

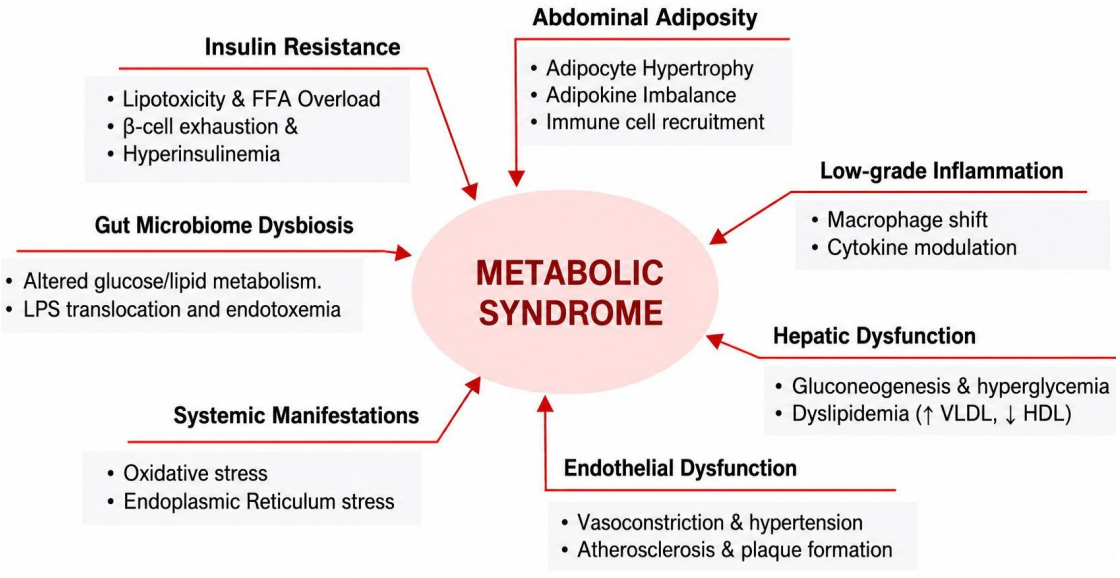


FIGURE 1.1: Pathophysiological mechanisms of metabolic syndrome [3].

1.1.1 Definition of Metabolic Syndrome

Metabolic syndrome (MetS) is a syndrome of interrelated metabolic abnormalities that tend to occur together in individuals and significantly increase the risk of cardiovascular disease and type 2 diabetes [4]. It is not a single disease, but rather a set of risk factors that, when present, indicate a much greater risk for severe downstream disease than any individual risk factor. From a clinical point of view, a person is considered to have metabolic syndrome when he or she has three or more of these five components, depending on the diagnostic definition used [5].

This diagnostic flexibility is due to the variety of professional organizations that have defined slightly different threshold values and various combinations of required thresholds; this may result in some variation in the definition of the condition between studies and populations. Whatever the criteria used, it is believed that most of the individual components of the syndrome are related to insulin resistance, and chronic low grade inflammation [6].

1.1.2 Historical Development of the Concept

The concept that multiple metabolic abnormalities occur in the same person is more than 30 years old prior to the introduction of the term "metabolic syndrome". This clustering was first described in early clinical observations, which go back almost a century, but it was the Banting Lecture by Gerald Reaven, in 1988, that put the focus firmly on insulin resistance in the process of establishing type 2 diabetes mellitus, and then on cardiovascular disease [6].

Reaven coined this combination "syndrome X," suggesting that the underlying cause of all these conditions is insulin resistance that leads to hyperinsulinemia, dysglycemia, high triglycerides levels, low HDL cholesterol and elevated blood pressure [7].

Reaven's original formulating did not include obesity as part of this definition. Since he could find non-obese people with insulin resistance and obese people without insulin resistance, he was unable to conclude that obesity was the common factor of the syndrome [7]. The term "syndrome X" was also changed to "metabolic syndrome X" to prevent confusion with the already known "syndrome X" in cardiology and the term "Reaven syndrome" was added to help identify the syndrome as a distinct clinical entity from the already known "syndrome X" in cardiology [6].

At the same time, the phrase "insulin resistance syndrome" also became popular in the literature, reflecting the importance Reaven attributed to insulin action. The concept progressed throughout the 1990s and 2000s, as international organizations grappled to put Reaven's theory into practice for clinical criteria.

Importantly, as other authors have noted, there have been instances of the concept of metabolic syndrome being confused with the diagnostic tools created later in the history of metabolic syndrome [7]. The widespread use of the term "metabolic syndrome" (not "syndrome X") coincided with a surge in research interest, culminating in the consensus that the group of abnormalities that Reaven had described should have a unified, clinical label [7].

1.1.3 Diagnostic Criteria

Metabolic syndrome is not a disease per se, but rather a cluster of risk factors, and various organisations have offered their own models over the years, with MetS having several different definitions such as the WHO, the NCEP-ATP III and the International Diabetes Federation (IDF) [8]. While these five components are shared, the way they are weighted and the threshold for diagnosis are different, leading to measurable differences in diagnosis across studies [9].

The WHO definition, which was originally formulated in the context of Reaven's model of insulin-resistance, mandates the presence of insulin-resistance or diabetes as a required component, in addition to any two of the other abnormalities including obesity, dyslipidaemia, hypertension or microalbuminuria [8]. The NCEP ATP III criteria adopt a different approach: the insulin resistance is not a requirement of metabolic syndrome in this case and all five components are equally weighted, and someone is considered to have metabolic syndrome if they have three of these five [5]. In the early 2000s, this change from the single mandatory criterion allowed for greater applicability of the ATP III in routine clinical and epidemiological practice, leading to its increased implementation.

IDF criteria were developed a little later and aimed to standardize the ATP III approach worldwide, by adding central obesity as a prerequisite (assessed by waist circumference) plus any two of the other 4 abnormalities [9]. One of the important innovations of the IDF definition was the use of ethnicity-specific waist circumference (WC) cut-offs based on evidence that the association between abdominal adiposity and cardiometabolic risk is different across ethnic groups.

WHO, ATP III and IDF screening tools have been applied to the same population groups and compared with each other in various studies to date, and consistently have yielded different prevalence estimates and have demonstrated only moderate agreement between the three, with the latter showing in southeast Nigeria a ranking of significant performance of harmonized, IDF, NCEP-ATP III and WHO [9]. This was a constant source of comment on the subject and eventually inspired the drive to harmonise the two definitions to a common standard.

1.2 Global Burden of Metabolic Syndrome

1.2.1 Worldwide Prevalence

The global prevalence of metabolic syndrome has shown a substantial and continuous increase over recent decades. Epidemiological evidence indicates that prevalence has risen markedly between 2000 and 2023, reaching approximately 31.0% among women and 25.7% among men worldwide. In absolute terms, this corresponds to an estimated 1.54 billion adults affected globally in 2023. The burden of MetS varies across regions, age groups, and socioeconomic strata, with higher prevalence observed in urbanized and higher-income populations, as well as with advancing age. Furthermore, studies also report overall global prevalence estimates around 25%, with increasing trends particularly in developing countries, largely attributed to lifestyle transitions and urbanization. These patterns underscore the expanding epidemiological burden and the need for targeted interventions [10].

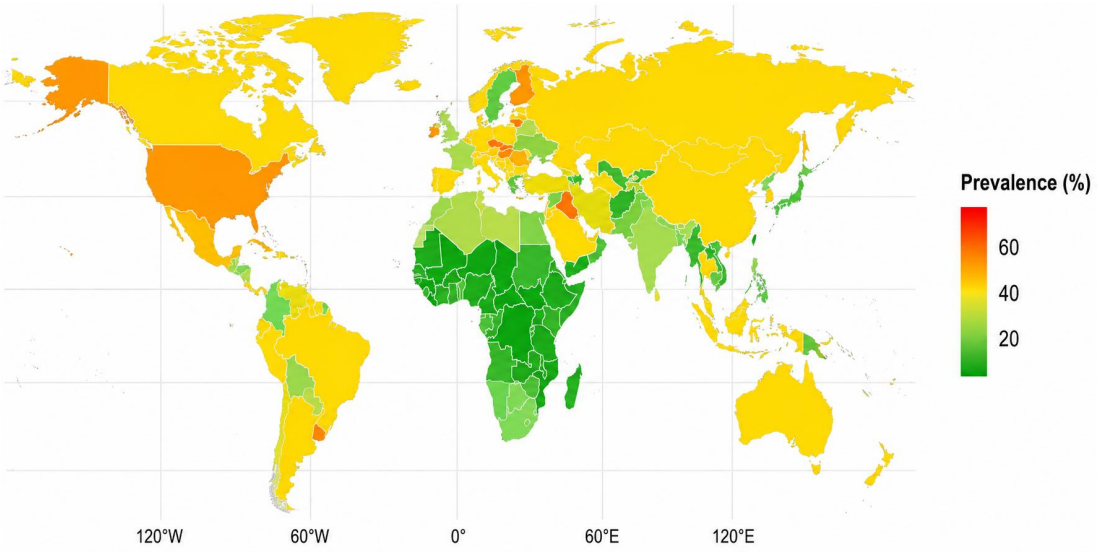


FIGURE 1.2: Global map of MetS prevalence in 2023 among men [11].

Studies shows between 2000 and 2023, the global prevalence of metabolic syndrome experienced a striking and widespread increase, rising from 11.9% to 28.4% across the adult population. This surge has resulted in an estimated 1.54 billion adults living with the condition worldwide as of 2023. The rising burden exhibits significant demographic and geographic disparities, with prevalence consistently

higher among women (31.0%) compared to men (25.7%), and escalating sharply with advancing age across almost all regions [12].

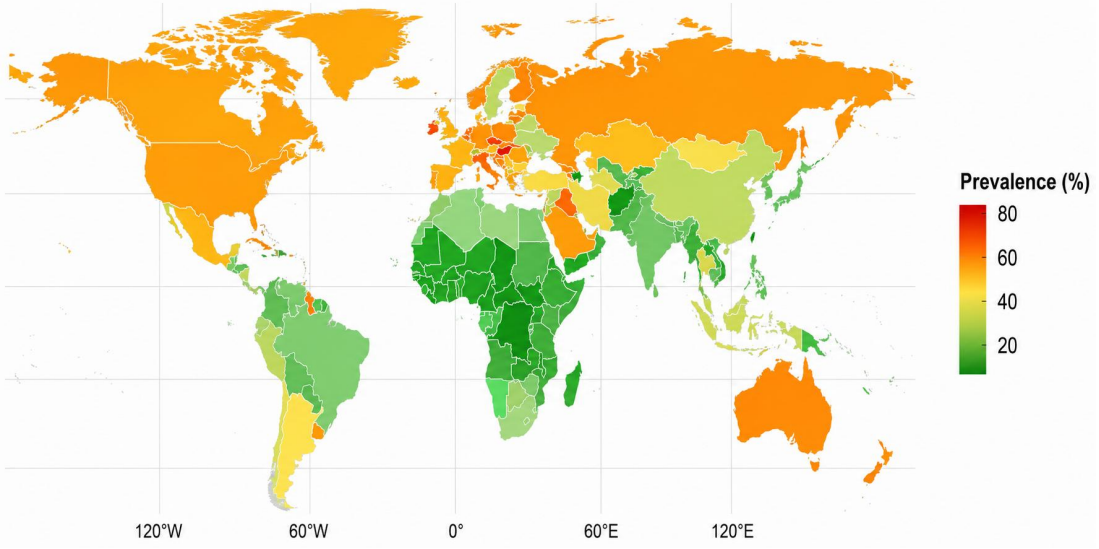


FIGURE 1.3: Prevalence of metabolic syndrome among men and women by country in different regions [11].

Furthermore, this upward trajectory is closely linked to socioeconomic and environmental shifts; prevalence increases alongside national income and urbanization levels, reflecting the detrimental impact of aging populations, westernized diets, and increasingly sedentary lifestyles. Notably, the prevalence of metabolic syndrome has increased in 196 countries and territories, underscoring its status as a pervasive and expanding global health crisis that requires urgent attention [13].

1.2.2 Prevalence in Asia

Asia has a significant burden of metabolic syndrome. According to a recent global modelling analysis [14] the prevalence of global adults with HIV increased from 14.7% in 2000 to 31.0% in 2023 and an estimated 1.54 billion adults were affected globally during that period. Decomposing this global burden, it is found that East Asia and South Asia were the top source of this burden both for men and women in 2023 [14], driven by the combined impacts of rapid urbanization, dietary transition and increasing levels of obesity across the continent. Figures from individual Asian countries are only slightly higher than or lower than the world average, with a

national study in Malaysia reporting that around a third (31.4%) of the general adult population had metabolic syndrome [15].

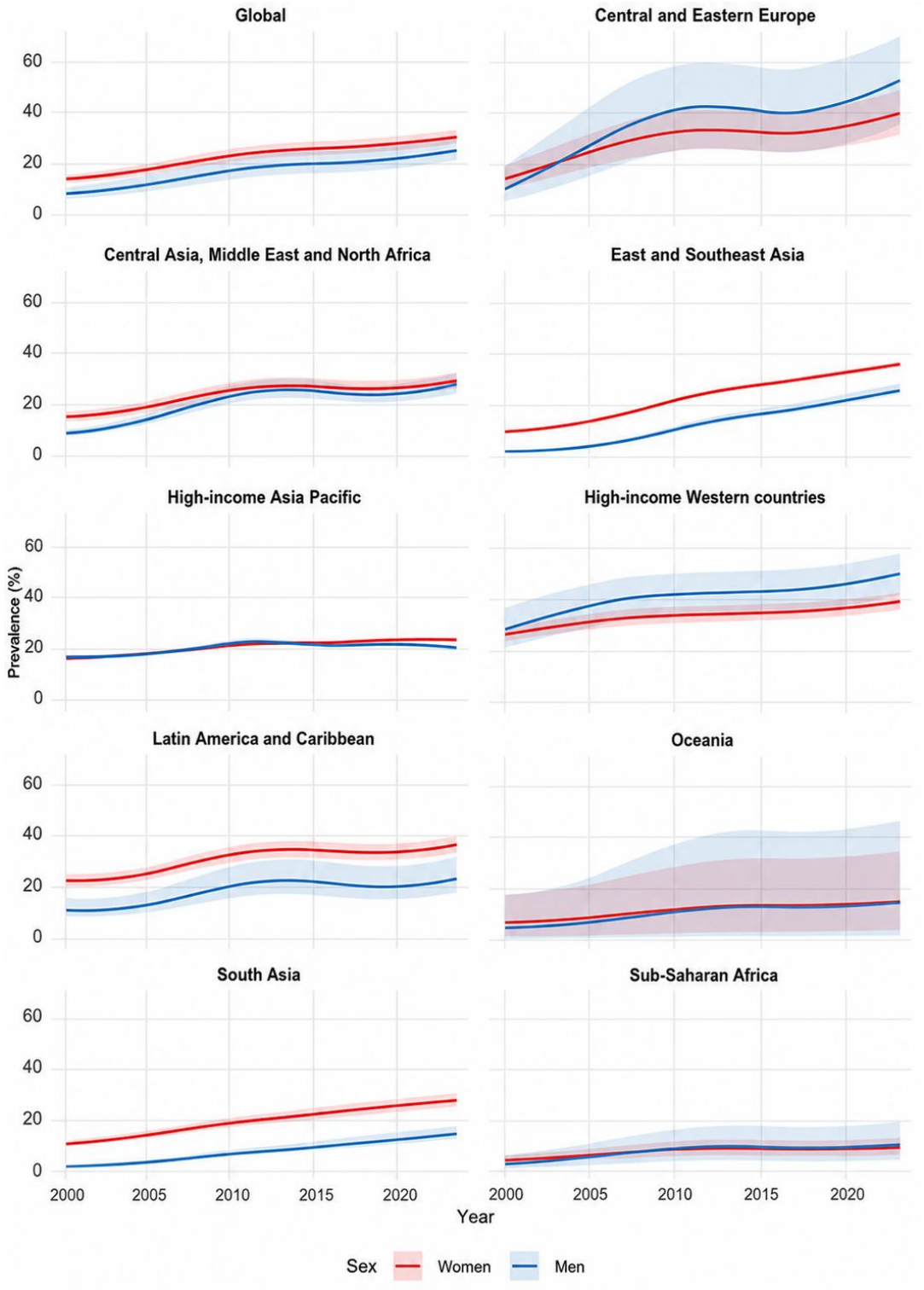


FIGURE 1.4: Global and regional temporal trends of metabolic syndrome prevalence by sex, 2000–2023 [11].

The risk seems to be higher among the South Asian populations compared to

other ethnic groups. In the various cross-sectional studies of metabolic syndrome (MS) prevalence among SA communities residing in countries outside their home continent, various ranges have been observed (27–47%) with a higher prevalence compared to other communities in the host country [16]. This is not only true of South Asians but is a pattern seen in many other ethnic groups, suggesting that many South Asian ethnic groups have a strong genetic and lifestyle-associated predisposition to central obesity and insulin resistance, two mechanistic aspects of the syndrome.

1.2.3 Prevalence in Pakistan

The situation in Pakistan is worse than the rest of the region. The prevalence of metabolic syndrome in the apparently healthy adult population was 28.8%, highest in a sub-urban village of Punjab (68%) and in province of Sindh (63.7%) and varied by diagnostic criteria with 33.2% using the IDF criteria and 23.9% using the NCEP criteria [13].

This is further supported by more recent data from the city level: A survey of the community in Karachi in 2022 found that prevalence was highest with the modified NCEP-ATP III definition (33.9%), followed by the IDF definition (32.2%) and the standard NCEP-ATP III definition (22.4%) [8]. Overall, these results suggest that in Pakistan, the prevalence of metabolic syndrome in adults may range from 25 to 28%, depending on the criteria used, indicating the magnitude of the cardiometabolic risk burden in the country.

1.2.4 Economic and Healthcare Burden

In addition to the clinical effects, metabolic syndrome is a significant and ever increasing economic burden in healthcare systems, due primarily to the downstream conditions it creates. One of the main cardiometabolic outcomes of the syndrome is diabetes, which is estimated to have consumed 12% of the world's health budget

in 2024 for an estimated cost of 1.015 trillion USD, an increase of 338% compared with the 17 years prior [17].

In a similar vein, the global economic burden of obesity is on a similar trajectory: global economic costs of overweight and obesity were estimated at 1.96 trillion US dollars in 2020, with expected losses approaching 4.32 trillion US dollars annually by 2035 when the economic costs of obesity includes treatment costs and lost productivity due to absenteeism, presenteeism and premature death [18]. As these figures demonstrate, the economic burden of each of the individual components is independent, and compounding, which is exactly what metabolic syndrome describes: that the more that occur, the more the economic burden on the patient increases.

A nine-year longitudinal study of linked national health insurance data at the individual level revealed that annual healthcare expenses of people with metabolic syndrome (MetS) significantly increased as they developed more complications, with morbid obesity alone increasing these by 58% and comorbid depression by up to 108% [19]. The cost attributable to microvascular complications (nephropathy, neuropathy, and retinopathy) was relatively small (24 to 27%) but significant compared to those without microvascular complications [19]. This finding suggests that the healthcare cost of metabolic syndrome is incremental and not a static cost, because each extra risk factor or complication of metabolic syndrome increases the cost of treating the syndrome.

This has seen the rising cost curve because of the scale of the syndrome. According to a global meta-analysis of 28+ million people, the prevalence of metabolic syndrome ranged from 12.5% to 31.4% depending on the definition used, with the Americas and Eastern Mediterranean experiencing the highest burden of metabolic syndrome [20]. As this is a significant proportion of the global population with the syndrome, health systems are presented with a correspondingly large number of patients who are at risk of the expensive downstream effects outlined above. Therefore, it is more cost effective to identify and prevent metabolic syndrome rather than treating downstream conditions [19].

1.3 Risk Factors of Metabolic Syndrome

1.3.1 Genetic Factors

Biomarkers play a critical role in the identification, diagnosis, and monitoring of metabolic syndrome. In this context, biomarkers are measurable biological indicators that reflect physiological and pathological processes associated with MetS. These include metabolic biomarkers such as elevated glycated hemoglobin (HbA1c), triglycerides, and fasting glucose levels; inflammatory markers such as C-reactive protein (CRP); and adipokines such as leptin and adiponectin. Additionally, immune-related markers, including increased CD3(+), CD4(+) T cells, and genetic factors such as APOC3 gene variants, are also implicated in disease susceptibility and progression. These biomarkers are essential not only for early detection but also for understanding disease mechanisms and evaluating therapeutic responses, thereby facilitating precision-based clinical management [21].

Metabolic syndrome is strongly associated with a wide spectrum of related disorders, most notably cardiovascular disease, type 2 diabetes mellitus, and metabolic dysfunction-associated liver disease, among others. The coexistence of its components amplifies the risk of complications such as atherosclerosis, hypertension-induced cardiac damage, and microvascular dysfunction. Additionally, conditions such as chronic kidney disease and polycystic ovary syndrome are also linked to MetS due to shared underlying mechanisms like insulin resistance and inflammation. The clinical and public health importance of metabolic syndrome lies in its role as an early indicator of these severe chronic diseases, making it a crucial target for early screening, prevention, and intervention strategies. Its growing prevalence and associated healthcare burden highlight the necessity for integrated approaches to mitigate its impact on global health systems [12].

Although extensive research has been conducted on biomarkers linked to metabolic syndrome, a major challenge remains in the dispersed nature of this information across the scientific literature. Biomarker-related findings are often published in

separate studies without adequate integration across molecular, clinical, and phenotypic levels. This lack of cohesion limits the ability to generate a holistic understanding and reduces the comparability of results across different studies. In addition, the absence of a centralized and systematically curated platform makes it difficult for researchers and clinicians to efficiently access and synthesize relevant information. Variations in terminology, measurement units, and reporting formats further contribute to inconsistencies, affecting reproducibility and limiting meta-analytical applications. As a result, these issues create significant obstacles in translating research findings into clinical practice and slow progress in improving diagnosis and management of metabolic syndrome [23].

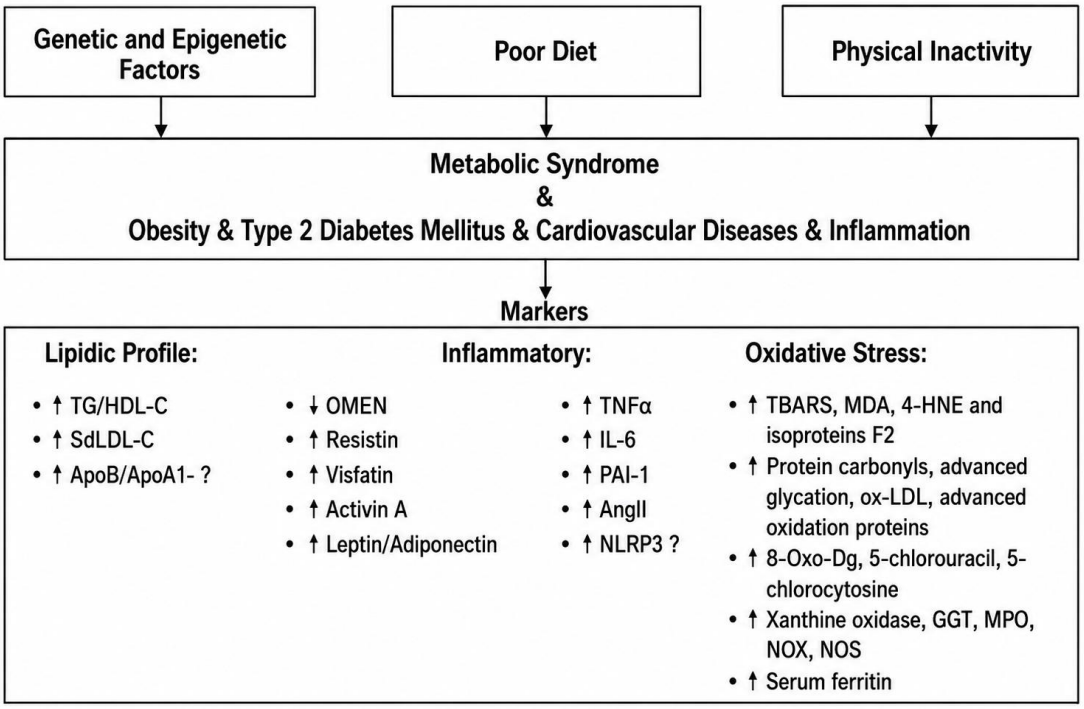


FIGURE 1.5: The main causes that can lead to MetS and its associated diseases [22].

To address these challenges, this study introduces the MetSMarker tool, designed as an integrated and well-structured platform for metabolic syndrome biomarker data. The main objective of this tool is to bring together scattered information from multiple scientific publications into a single, organized system that supports easy access and meaningful comparison. By applying standardized annotation methods and uniform data processing approaches, the platform ensures consistency and improves data reliability.

Beyond simple data collection, MetSMarker also functions as an exploratory system that assists in interpreting biomarker relationships and supports both research activities and clinical decision-making. Its design aligns with the principles of precision medicine by enabling improved risk assessment, diagnostic support, and monitoring of disease progression, thereby enhancing the practical application of biomarker knowledge.

This study focuses on identifying, extracting, and organizing a broad range of metabolic syndrome-related biomarkers, including molecular, biochemical, and clinically relevant indicators reported in peer-reviewed scientific studies. The data included in the platform are derived exclusively from published literature to ensure scientific validity and reliability.

The work also involves the development of the MetSMarker web-based platform, which provides structured data storage, standardized representation, and an intuitive interface for efficient exploration and analysis of biomarker information. While the study offers a comprehensive integration framework, it is limited by dependence on existing published data and the variability in study quality across sources. Nevertheless, it provides a strong foundation for future enhancements, including expansion of datasets and incorporation of more advanced analytical capabilities.

1.3.2 Lifestyle Factors

One factor that has been consistently and strongly reported to be a driver of metabolic syndrome is modifiable lifestyle behaviors. A clear dose-response relationship exists for sedentary time and risk of metabolic syndrome: a systematic review and meta-analysis showed that moderate (median 4.1 hours/day) and high sedentary time (median 7.3 hours/day) were both associated with increased risk of metabolic syndrome [24]. This phenomenon is associated with the direct effect of prolonged sitting on decreasing skeletal muscle glucose uptake and lipoprotein lipase activity, which are two characteristics directly linked to central obesity, dyslipidemia and insulin resistance, respectively, in the syndrome.

Dietary pattern is also a consistent contributor. A study of dietary patterns and physical activity combined showed that those with a prudent, active lifestyle pattern did not have significantly higher odds of any of the components of metabolic syndrome than those with a non-prudent, low activity pattern, with the exception of 43% lower odds of central obesity and 48% lower odds of hypertension [25].

In comparison, the highest prevalence of metabolic syndrome out of the three behavioral clusters was with the Western dietary pattern (78.8%), followed by the prudent-active cluster (9%) and the balanced cluster (12.2%) [25].

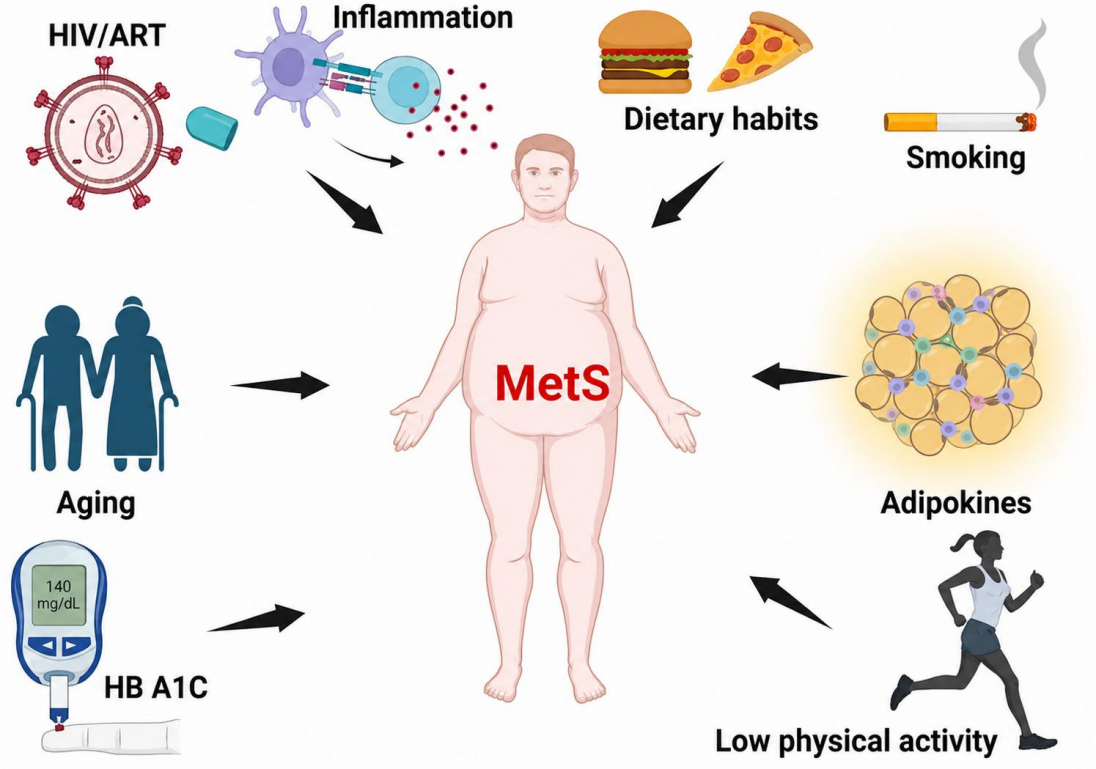


FIGURE 1.6: Factors that increase the risk of metabolic syndrome [21].

Other behaviors contribute to this risk, in addition to diet and activity. More than 4 hours of TV viewing daily was associated with higher central adiposity, lower HDL cholesterol, and higher triglycerides and glucose levels among genetically susceptible young Saudi adults [26].

These exposures are often combined in the same person; therefore, studying lifestyle patterns is becoming more common than studying each behavior individually, for the purpose of determining which are at greatest risk when combined in a person.

1.3.3 Environmental Factors

In addition to personal behaviour, ambient exposures are now recognized as an important and independent risk factor for metabolic syndrome. The results are from a nationwide study in China that showed that exposure to particulate matter was significantly associated with central obesity, elevated blood pressure, elevated fasting glucose and low HDL cholesterol, and for each 10 $\mu g/m^3$ increase in NO_2, the adjusted odds of metabolic syndrome increased by over twofold [27].

It is an example that shows that air pollution is not only linked to the syndrome in general but that it is also, independently, associated with multiple components of the diagnosis of the syndrome. Mechanisms of this relationship are beginning to be understood.

A study of environmental pollution and metabolic disorders revealed that fast industrialisation has created excessive amounts of endocrine-disrupting substances such as persistent organic pollutants (POPs) that disrupt adipogenesis; heavy metals have also been shown to directly affect the function of pancreatic beta-cells and glucose regulation [28].

These compounds work on the same hormonal and metabolic mechanisms involved in insulin resistance, so it is believed that their effect is additive rather than a completely independent pathway to insulin resistance. Urbanization seems to be another environmental factor, apart from pollution. Urban environmental pressures on the endocrine system have been reviewed, and urban-specific stressors such as pollution, artificial light, noise, and sedentary lifestyle, are all acting synergistically on endocrine function [29].

1.3.4 Age and Gender

The prevalence of metabolic syndrome increases steadily with age in practically all populations examined, although the trends with sex are more complex and change over the lifespan. A global modelling analysis based on data to 2023 indicated that there was an age gradient for both sexes, with higher prevalence at higher

urbanization levels and among individuals with higher income, and that women had a higher prevalence globally than men in 2023 [14]. This higher prevalence among females, however, is a physiological "reversal" when young; men vs women of the same age have a higher prevalence of the syndrome, with this pattern closely associated with the protective metabolic effects of estrogen before menopause.

This is a turning point in menopause. A large community-based study of Asian women reported that postmenopausal women were at about 2 times risk for metabolic syndrome than premenopausal women (30% vs. 14%), had significantly higher odds after adjusting for age, and that surgical menopause was associated with higher odds than natural menopause [30]. This was confirmed by a systematic review and meta-analysis in Iranian postmenopausal women reporting a pooled prevalence of MS around 58%, which is a sharp increase after the cessation of estrogen's protective effect [31].

These results collectively indicate that the idea of age and sex as static risk groups is too simple, and that there is a dynamic interplay between the two concepts over the life course. Since the postmenopausal period carries a disproportionate burden of the female metabolic syndrome and the experience of the menopausal transition is unique, there is a growing interest from the clinician and public health research perspective to identify the menopausal transition as a window for screening and preventive interventions.

1.4 Clinical Importance of Early Diagnosis

Metabolic syndrome (MetS) is a progressive cardiometabolic syndrome that is associated with central obesity, insulin resistance, dyslipidemia, hypertension, and impaired glucose metabolism. This abnormality may develop slowly and not cause any symptoms for years, which means the syndrome is often underdiagnosed until irreversible organ damage or serious heart events take place. Early diagnosis allows the identification of those who have an increased cardiometabolic risk prior to the manifestation of the disease and allows the implementation of lifestyle changes

and pharmacological treatment in a timely manner. These preventive measures have been found to slow disease progression, enhance metabolic control and lower long-term health care costs due to diabetes and cardiovascular diseases [32, 33].

The importance of early diagnosis in regards to the clinical significance is not limited to the components of metabolic syndrome. Early detection enables full assessment, ongoing monitoring and tailored therapeutic approaches for multiple metabolic abnormalities at once. There is also emerging data indicating that using novel biomarkers and imaging modalities as part of routine clinical assessment can facilitate early detection and risk stratification, potentially preventing complications before they occur since tissue can then still be salvaged [34, 35].

1.4.1 Disease Progression

Metabolic syndrome is a complex condition caused by low-grade inflammation, unhealthy eating habits, sedentary lifestyle, and genetics. Insulin resistance and central adiposity are the earliest changes, followed by hyperinsulinemia, dyslipidaemia, endothelial dysfunction, and high blood pressure. If left untreated, these abnormalities continue to cause the accumulation of progressive metabolic impairment, which in turn promotes the onset of type 2 diabetes mellitus (T2DM), atherosclerosis and cardiovascular disease. Importantly, MetS is a progressive disease that slowly develops, providing valuable opportunities for intervention before irreversible metabolic dysfunctions [32].

Along with the secretion of pro-inflammatory cytokines, including tumor necrosis factor-α and interleukin-6 (IL-6), and through the imbalance of adipokines, insulin resistance leads to chronic inflammation, which in turn drives metabolic syndrome.Insulin resistance also contributes to metabolic syndrome by inducing pro-inflammatory cytokines (such as tumor necrosis factor-α (TNF-α) and interleukin-6 (IL-6)), as well as the imbalance of adipokines and oxidative stress.

These processes lead to endothelial injury, vascular remodeling, hepatic steatosis and pancreatic β-cell dysfunction, and finally to the development of diabetes and

cardiovascular complications. The timely diagnosis of these metabolic derangements allows interventions such as dietary changes, regular exercise, weight loss, and specific pharmacotherapy to stop or at least slow the progress of the disease, leading to a dramatic improvement in long-term outcomes [34, 35].

1.4.2 Associated Complications

The risk of many organ system related chronic diseases is greatly elevated when metabolic syndrome is not diagnosed in the initial phase. Cardiovascular disease continues to be the major complication, and the risk of having a myocardial infarction, stroke, heart failure and cardiovascular mortality is significantly elevated in those who have been diagnosed with MetS.

Moreover, a hyperinsulinemic state caused by insulin resistance may lead to progressive pancreatic β-cell failure, leading to type 2 diabetes mellitus. Other identified issues are non-alcoholic fatty liver disease (NAFLD), chronic kidney disease, obstructive sleep apnea and some obesity-associated cancers [32, 35].

Metabolic syndrome also promotes systemic inflammatory response and endothelial dysfunction, thereby worsening the damage to the vascular units and the impairment of the organs. Some evidence is now emerging that metabolic abnormalities cause cognitive decline, poor quality of life and increased all-cause mortality with prolonged exposure. Hence, early detection of metabolic syndrome is crucial to lower morbidity and better prognosis in the long-term [34].

1.4.3 Importance of Early Detection

Metabolic syndrome can be prevented if all the pathological alterations that characterize the syndrome are reversible if detected early enough and treated appropriately, which is the basis of preventive medicine. The identification of at-risk individuals enables health care providers to make evidence-based decisions regarding dietary modification, increased physical activity, smoking cessation, weight management and/or pharmacologic treatment if necessary. These interventions

have a significant effect on insulin resistance, lipid metabolism, blood pressure and prevent or delay type 2 diabetes and cardiovascular disease [32].

In recent years, biomarker research has also emphasized the importance of the early detection. In addition to classic criteria, new biomarkers may also help determine metabolic disturbances even before traditional clinical symptoms appear, including adiponectin, leptin, high-sensitivity C-reactive protein (hs-CRP), uric acid, inflammatory cytokines, and markers of oxidative stress. Conventional metabolic parameters may be used in concert with novel biomarkers to enhance diagnostic sensitivity, allow an individualized approach to risk prediction, and aid a precision medicine approach to the management of metabolic syndrome [35, 36].

1.5 Role of Biomarkers in Metabolic Syndrome

Metabolic syndrome (MetS) is a condition of a complex nature with multiple interconnected metabolic abnormalities such as insulin resistance, central obesity, dyslipidaemia, hypertension and impaired glucose metabolism. Because of multifactorial pathophysiology, no one laboratory parameter is able to accurately reflect onset, progression, and clinical outcome. Biomarkers have thus become useful tools to detect metabolic changes linked to MetS, which offer objective measures of metabolic dysfunctions that can be used in addition to the standard clinical assessment.

Biomarkers not only improve diagnostic accuracy, they are also used to improve the risk stratification, monitoring disease progression and to provide a basis for tailored therapeutic interventions. In recent years, new molecular biology and metabolomics have been developed, and many of these factors have shown strong correlations with cardiometabolic risk [6], such as the inflammatory mediators, adipokines, oxidative stress markers, endothelial dysfunction markers and circulating microRNAs.

The rising interest in biomarker research is due to the fact that the metabolic changes associated with MetS may occur years before clinical symptoms appear.

Traditional parameters such as fasting plasma glucose, serum triglycerides, high-density lipoprotein cholesterol (HDL-C), waist circumference and blood pressure are still the basics for diagnosis. These indicators mostly detect the earliest pathological changes, however. New biomarkers could be used to identify subclinical inflammation, adipose tissue dysfunction, insulin resistance, and endothelial injury before currently accepted thresholds are met, thus providing better chances for early intervention and personalised management [37].

1.5.1 Diagnostic Biomarkers

Diagnostic biomarkers are biological markers that can be used to detect people who have metabolic syndrome or are at high risk to develop the syndrome before they have irreversible metabolic damage. The current diagnostic practice is mainly based on the traditional metabolic parameters such as fasting plasma glucose, triglyceride, HDL cholesterol, blood pressure and central obesity parameters. These biomarkers are still clinically useful due to their ease and standardization, but offer little information on the molecular and inflammatory mechanisms that initiate disease. This has led to extensive studies of seeking complementary markers that would enhance the diagnostic sensitivity and specificity [6].

High-sensitivity C-reactive protein (hs-CRP) is one of the most widely studied diagnostic markers, as it is an acute-phase protein that is indicative of chronic low-grade inflammation that is associated with MetS. An increase in hs-CRP levels has repeatedly been linked to obesity, insulin resistance, endothelial dysfunction and cardiovascular risk. Adipokines also have a significant role in the diagnosis. This suggests that adiponectin is reduced and leptin is elevated, which are associated with poor functioning of adipose tissue to be associated with insulin resistance and abdominal fatness. In addition, the LpA ratio has proved superior to leptin and adiponectin in isolation as a biomarker for diagnosis based on the balance between pro-inflammatory and insulin-sensitizing pathways [37].

Other markers of diagnostic value include inflammatory cytokines like interleukin-6 (IL-6) and tumour necrosis factor alpha (TNF-α), oxidative stress markers,

gamma-glutamyl transferase (GGT), serum uric acid, ferritin and fetuin-A. These biomarkers include different components of metabolic dysfunction such as hepatic steatosis, oxidative injury, chronic inflammation and endothelial injury. While most are still in the investigational stage, there is growing evidence that the combination of conventional metabolic markers and novel biochemical markers may significantly advance early diagnosis and detection of those who need preventive intervention before the onset of metabolic disease [37].

1.5.2 Prognostic Biomarkers

Prognostic biomarkers are markers that give information about the progress of the disease, the possible complications and/or long-term clinical results after the diagnosis of metabolic syndrome. Biomarkers with predictive value for the progression of disease are of great clinical value, given that MetS is a strong risk factor for type 2 diabetes mellitus, cardiovascular disease, chronic kidney disease and non-alcoholic fatty liver disease (NAFLD). Prognostic biomarkers can help identify low and high-risk patients and facilitate earlier initiation of intensive lifestyle and pharmacological interventions [6].

Inflammatory markers continue to be some of the best prognostic factors in MetS. High-levels of hs-CRP, IL-6, TNF-α, fibrinogen and plasminogen activator inhibitor 1 (PAI-1) are linked to rapid atherosclerosis, endothelial dysfunction and higher rates of cardiovascular events. Similarly, low levels of adiponectin are linked to deteriorated insulin resistance, increased risk of type 2 diabetes and poor cardiovascular (CV) outcomes, while high levels of leptin are linked to inflammation and metabolic deterioration in obesity. Longitudinal studies have shown that people with long-lasting elevated inflammatory and adipokine levels have much higher cardiometabolic risk than people with normal levels of these biomarkers [6].

A number of biochemical markers have also been found to have a prognostic value. Serum uric acid, GGT, liver enzymes, and ferritin are further associated with liver fat content, oxidative stress, increasing insulin resistance and are predictive of the

progression to NAFLD and cardiovascular disease. Recently, circulating microRNAs are shown as potential markers for prognosis due to the fact that microRNAs regulate genes that are associated with lipid metabolism, glucose homeostasis, adipogenesis and inflammatory signalling. While additional validation work is needed before clinical use, these molecular biomarkers could have a significant impact on long-term risk prediction and personalised patient monitoring [38].

1.5.3 Predictive Biomarkers

Predictive biomarkers are biological markers that can be used to predict a person's risk of developing metabolic syndrome and/or response to preventive or therapeutic measures. Predictive biomarkers, unlike diagnostic markers, allow assessing risk prior to the onset of clinical symptoms. As their identification becomes more critical, metabolic syndrome tends to occur over a number of years without being noticed, thus providing time to intervene and prevent disease. With the development of genomics, proteomics, metabolomics, and molecular medicine, the identification of predictive biomarkers has increased rapidly, with the ability to identify high-risk individuals, more accurately than just clinical measurements [39].

The best potential markers are adiponectin, leptin, hs-CRP, insulin, HOMA-IR, and the leptin/adiponectin ratio. Low levels of adiponectin and high levels of inflammatory markers have been shown to be associated with a significantly higher risk of metabolic syndrome in prospective cohort studies. Likewise, high hs-CRP is a risk factor for insulin resistance, type 2 diabetes and cardiovascular disease, underscoring the important role of chronic inflammation in metabolic dysfunction. Inflammatory biomarkers could be used together with anthropometric and biochemical measurements to enhance the prediction models and to identify susceptible individuals early [37]

Circulating microRNAs are one of the most dynamic new molecular biomarkers in the field of predictive medicine. MicroRNAs control many metabolic pathways that are implicated in glucose metabolism, differentiation of adipocytes, insulin signalling, oxidative stress and inflammatory responses. Multiple miRNAs

have been identified through systematic reviews that are consistently linked with metabolic syndrome and its components. Though these biomarkers, are still not included into our daily clinical practice, yet they still demonstrate a considerable potential for the early risk prediction as well as personalized therapeutic decision-making. Future integration of the conventional biochemical markers along with the genomic and molecular biomarkers, is believed to improve the precision medicine approaches for not only preventing metabolic syndrome but also its associated complications [38].

1.6 Need for an Integrated Biomarker Platform

1.6.1 Challenges in Existing Literature

Although significant advances have been made in the identification of biomarkers linked to metabolic syndrome, important limitations remain evident in the current scientific literature. A major gap lies in the lack of integration across studies, where biomarker data are often reported in isolation without linking molecular, clinical, and phenotypic dimensions. This fragmentation restricts comprehensive understanding and limits cross-study comparability. Furthermore, accessibility remains a significant challenge, as biomarker information is dispersed across multiple publications and databases, making systematic retrieval and utilization both time-consuming and inefficient. Another key limitation is the absence of standardized data representation [40]. Variability in biomarker nomenclature, measurement units, and reporting formats introduces inconsistencies that hinder reproducibility and meta-analysis. Collectively, these issues underscore the need for a unified and systematically curated resource.

1.6.2 Importance of Data Integration

The significance of identifying and monitoring metabolic syndrome stems from its cumulative impact on disease progression and mortality, as it comprises a complex

group of interrelated conditions such as insulin resistance, central obesity, high blood pressure, and dyslipidaemia. Monitoring these co-occurring factors is clinically vital, as their clustering accelerates disease progression, complicates patient management, and confers a substantially higher cardiovascular risk than isolated individual components [41]. Metabolic syndrome functions as an important public health marker for detecting individuals who face a significantly increased risk of developing serious long-term outcomes, including type 2 diabetes, cardiovascular disease, chronic kidney disease, metabolic dysfunction-associated steatotic liver disease, multiple types of cancer, and early mortality. Given that these interconnected metabolic components are major contributors to global disability-adjusted life years and mortality, the sheer scale and complexity of the syndrome demand integrated clinical strategies and robust data tracking to effectively manage and mitigate its growing global burden [42].

1.6.3 Role of Bioinformatics Database

Bioinformatics databases are crucial for research on metabolic syndrome (MetS) as they serve as a centralized platform for the storage, organization, and analysis of vast amounts of biological and clinical data. These databases combine data on the genome, transcriptome level, protein, metabolites and clinical information, allowing the identification of genes, proteins, pathways and candidate biomarkers that are associated with the pathogenesis of MetS. [43] mentioned that reliable tools like the GWAS Catalog are useful for identifying genetic variants linked to metabolic disorders, and [44] noted that Human Metabolome Database (HMDB) offers extensive information about human metabolites and their functions, which is useful for discovering biomarkers and studying metabolic pathways. In a similar way, [45] stated that STRING can also be used to build protein–protein interaction networks, which can be useful for determining the importance of regulatory proteins and pathways involved in disease development. Furthermore, KEGG is a widely used pathway enrichment analysis tool and a functional interpretation of high throughput data (omics) [46]. Together, these bioinformatics databases have facilitated the discovery of novel diagnostic and prognostic biomarkers, enhanced

understanding of molecular mechanisms of metabolic syndrome, and advances in precision medicine and personalized therapeutic strategies.

1.7 Problem Statement

Three major issues plague the present literature on metabolic syndrome biomarkers, despite notable advancements in the field: Fragmented integration: Biomarker data are usually examined separately and infrequently connected to clinical or phenotypic context, which restricts thorough comprehension and cross-study comparability. Inadequate accessibility: Systematic retrieval and synthesis are time-consuming and ineffective since pertinent findings are dispersed throughout multiple publications and databases with varying breadth and organization. Inconsistent reporting formats, measurement units, and nomenclature [40] lower repeatability and impede meta-analysis. All of these limitations point to the need for a single, methodically curated database that combines data on metabolic syndrome biomarkers from molecular, clinical, and phenotypic perspectives.

1.8 Aims and Objectives

1.8.1 Aim

To integrate experimentally validated molecular biomarkers associated with metabolic syndrome through comprehensive manual curation, enabling biomarker characterization, interactive exploration, comparative analysis, and translational research by developing a web-based knowledge exploration platform.

1.8.2 Objectives

i. To systematically identify, extract, and manually curate experimentally validated molecular biomarkers associated with metabolic syndrome from peer-reviewed scientific literature.

ii. To standardize and annotate the curated molecular biomarkers by harmonizing, molecular classifications, standardized identifiers, and biological annotations to ensure data consistency, interoperability, and comparative analysis.

iii. To enable biomarker exploration, comparative analysis, and translational research toward precision medicine in metabolic syndrome through an accessible web-based knowledge exploration platform.

1.9 Scope and Significance

The systematic curation, categorization, and computational arrangement of empirically verified molecular biomarkers linked to metabolic syndrome as documented in peer-reviewed scientific publications is the exclusive focus of this project. As stated in Objectives 1 and 2, the scope covers the molecular properties, biological roles, experimental data, and diagnostic relevance of the four main biomarker categories: DNA-based, RNA-based, protein-based, and epigenetic. The study's scope includes designing and developing a web-based platform with three functional modules based on this curated dataset: a Browse module for organized retrieval and filtering of biomarker information, an Analyse module for comparative investigation of shared biological characteristics and functional relationships among biomarkers to aid in prioritization and hypothesis generation, and an Insights module for statistical summarization and visualization of biomarker distributions, classifications, and publication trends. In keeping with the main goal of the project, these elements work together to create an easily accessible platform that supports biomarker discovery and translational research in metabolic syndrome.

1.9.1 UN Sustainable Development Goals

The scope of this study aligns directly with several UN Sustainable Development Goals, situating the work within a broader global development agenda rather than as a purely academic exercise.

i. SDG 3 (Good Health and Well-being) is the most direct point of alignment. SDG target 3.4 specifically calls for reducing premature mortality from non-communicable diseases, including cardiovascular disease and diabetes, by one third by 2030 [47]. Because metabolic syndrome is a principal upstream driver of both conditions, a curated biomarker platform that supports earlier detection and more precise risk stratification contributes directly to the prevention and treatment efforts this target calls for.

ii. SDG 9 (Industry, Innovation, and Infrastructure) is reflected in the study's focus on creating a structured digital research infrastructure, the Browse, Insights, and Analyse modules, which did not previously exist in a consolidated form for metabolic syndrome biomarkers and strengthens scientific research infrastructure and capacity in the health informatics domain.

iii. SDG 17 (Partnerships for the Goals), particularly its emphasis on data availability and knowledge sharing, is reflected in the study's objective of integrating scattered, cross-institutional biomarker evidence into a single accessible, shareable resource, supporting the kind of open, collaborative research infrastructure this goal envisions.

The study presents itself as a translational contribution to both precision medicine research and more general global health and innovation aims by limiting its scope to curation, categorization, and platform creation while clearly connecting these outputs to SDG 3, 9, and 17.

1.9.2 Significance

In light of these limitations, the development of MetSMarker is both timely and necessary. The proposed platform addresses the fragmentation of existing knowledge by integrating biomarker data from diverse studies into a centralized and structured repository [48]. By applying standardized ontological annotations and harmonized data preprocessing protocols, MetSMarker ensures consistency and comparability across datasets. Additionally, the platform enhances accessibility

by providing a user-friendly interface for efficient biomarker exploration and retrieval. Unlike conventional approaches that focus solely on data collection, MetS-Marker incorporates knowledge exploration modules that enable functional interpretation, precision medicine insights, and multi-dimensional analysis. This integrative framework not only improves data usability but also facilitates the identification of clinically relevant biomarkers, thereby advancing diagnostic research in metabolic syndrome.

Chapter 2

Literature Review

Biomarker research has emerged as a cornerstone of modern biomedical science, offering critical insights into the diagnosis, prognosis, and mechanistic understanding of complex disorders such as metabolic syndrome (MetS). Over the last twenty years, considerable research has been directed toward discovering molecular, genetic, and clinical biomarkers that capture the complex nature of metabolic syndrome (MetS), a condition characterized by linked disorders such as insulin resistance, obesity, dyslipidemia, and hypertension. [1].

These biomarkers—ranging from circulating proteins and inflammatory mediators to genetic variants—serve as measurable indicators of disease onset and progression, enabling early detection and risk stratification. In the context of precision medicine, biomarker discovery plays a pivotal role in advancing personalized therapeutic strategies by linking biological signatures with clinical phenotypes. Consequently, the growing body of literature underscores the importance of systematically characterizing and integrating biomarker data to enhance diagnostic accuracy and improve clinical outcomes in metabolic syndrome [21].

Insulin resistance is considered the main mechanism linking the various parts of metabolic syndrome. In addition to insulin resistance, there is also a second pathobiological feature of metabolic syndrome, chronic low-grade inflammation, which is now known to be closely linked to the insulin resistance pathway and not a distinct entity [53].

2.1 Pathophysiology of Metabolic Syndrome

2.1.1 Insulin Resistance

MetS is a complex process involving a number of acquired and inherited components under the umbrella of insulin resistance and chronic low-grade inflammation [6]. Insulin normally increases glucose uptake in skeletal muscle, liver, and adipose tissue, and decreases output of glucose from the liver and lipolysis, and when these effects are lost the body responds by increasing insulin and, over time, leads to hyperglycemia, altered lipid handling, and vascular dysfunction [49].

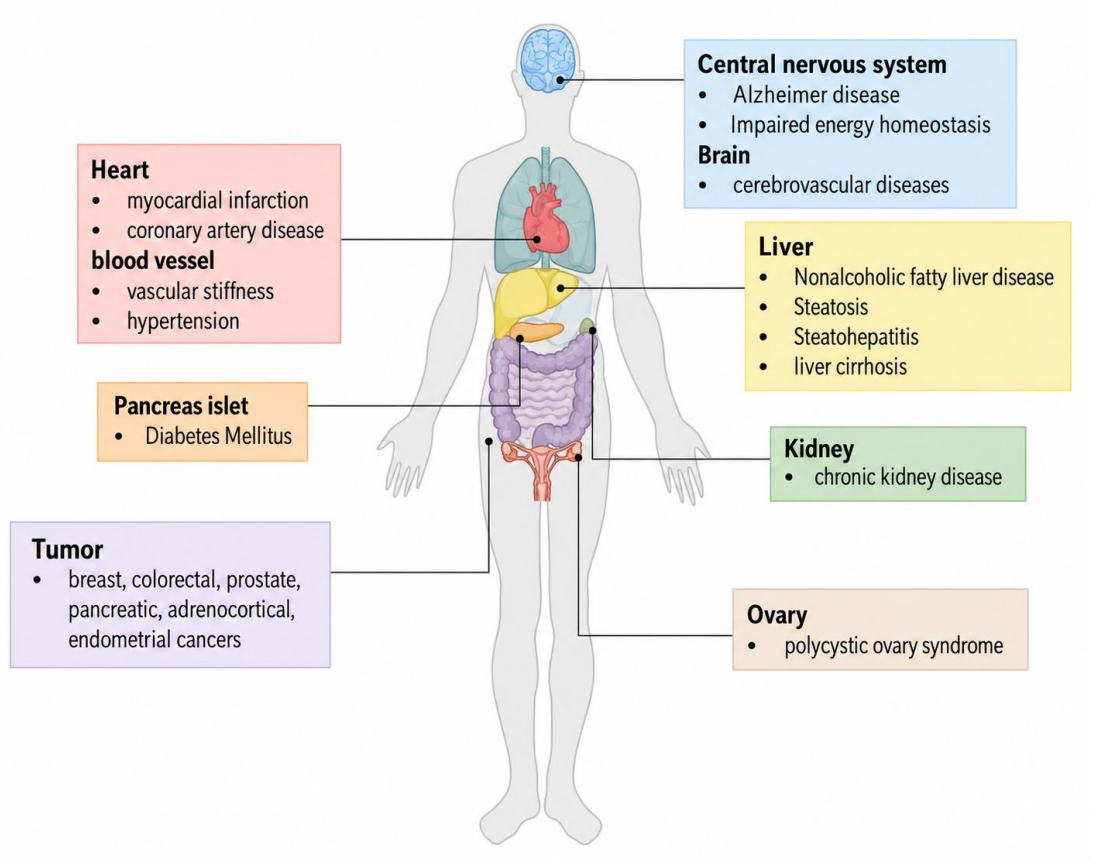


FIGURE 2.1: Chronic metabolic diseases may be induced by IR [51].

Now, lipid deposition in the wrong places is known to be among the fundamental causes of this resistance. Insulin resistance is a complex pathophysiological process that starts with a genetic susceptibility and is triggered by core metabolic disturbances, primarily lipid accumulation in tissue where lipids are not normally stored (such as liver and skeletal muscle), and is then gradually exacerbated by cell stress

and systemic inflammation [50]. In this ectopic fat deposition, insulin signaling pathway is directly affected and affects the translocation of glucose transporters to the cell membrane and affects the normal disposal of glucose.

The effects of such resistance are much wider-ranging than glucose control. Insulin resistance leads to liver fat accumulation by stimulating lipolysis, increasing free fatty acid flux, and decreasing fatty acid oxidation, and a metabolic syndrome state further reinforces this vicious cycle by promoting visceral adiposity, chronic inflammation and impaired lipid metabolism [52]. Insulin resistance is upstream of hyperglycemia, dyslipidemia and hypertension, so it is often considered the first abnormality of the natural history of metabolic syndrome, and a primary target for prevention before the other aspects of metabolic syndrome are fully developed.

2.1.2 Chronic Inflammation

Adipose tissue is no longer considered merely a storage depot for fat but rather an active endocrine organ that regulates adipose tissue inflammation via endocrine, paracrine, and autocrine effects of proteins secreted by adipose tissue, liver, and skeletal muscle, respectively, namely adipokines, hepatokines, and myokines [53]. In obesity, the secretions of adipose tissue tend to be clearly pro-inflammatory.

Obesity is linked to increased levels of pro-inflammatory adipokines (leptin, resistin and chemerin), and lower levels of anti-inflammatory adipokines (adiponectin), which results in a systemic low-grade inflammatory state [53]. This change is followed by changes within the tissue itself, including accumulation of macrophages and activation of NF-κB, JNK, and the NLRP3 inflammasome pathways, each of which directly impact local insulin signaling [53]. It is a serious inflammatory condition, and the inflammation is not limited to adipose tissue; it affects other areas of the body as well. Chronic low-grade inflammation of adipose tissue is linked to a variety of downstream metabolic disorders, and its complex interplay between pro-inflammatory and anti-inflammatory mechanisms of the immune response within adipose tissue is now recognized as a major contributor to the overall metabolic dysfunction observed in the syndrome [55].

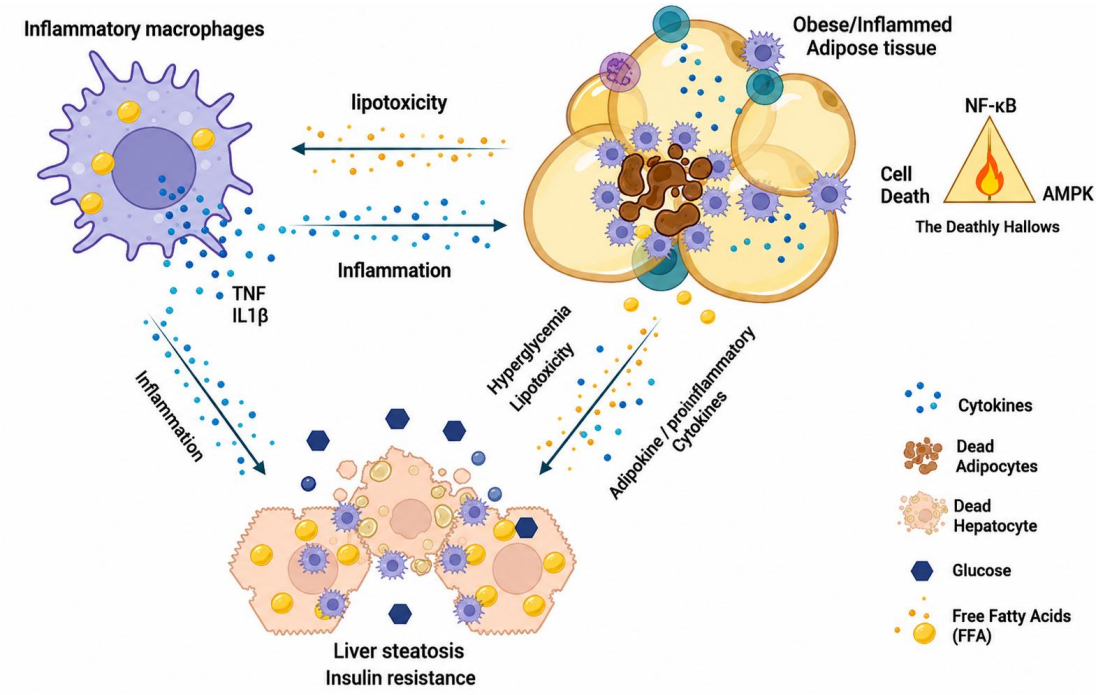


FIGURE 2.2: Chronic inflammation, Obesity and IR [54].

Given that this inflammation is a cause of insulin resistance and also contributes to it, these pathways are now being increasingly viewed as a single entity than as independent factors influencing the insulin resistance syndrome.

2.1.3 Oxidative Stress

The third mechanism that is orchestrated in metabolic syndrome is the oxidative stress, which is the balance between the generation of reactive oxygen species (ROS) and the body's ability to neutralize them. Both inflammation and oxidative stress favor the subsequent development of the different complications associated with the components of the syndrome: insulin resistance, hypertension, hyperlipidaemia, etc. [56].

The production of ROS is a result of the activation of enzymes in the cytosol, cell membrane and mitochondria and in the absence of sufficient antioxidant capacity, the ROS lead to oxidative stress and create intracellular damage [56]. There are two enzymatic sources of ROS that are especially involved in this process. The interplay between NADPH oxidases and mitochondria is the primary mechanism

responsible for the generation of ROS that exacerbates metabolic syndrome-related chronic inflammation, adipocyte proliferation and insulin resistance, and hence, the link between oxidative stress and other key pathways of metabolic syndrome [57]. This is a dual-source ROS production, which enhances inflammatory signaling mechanisms instead of going down a different pathway.

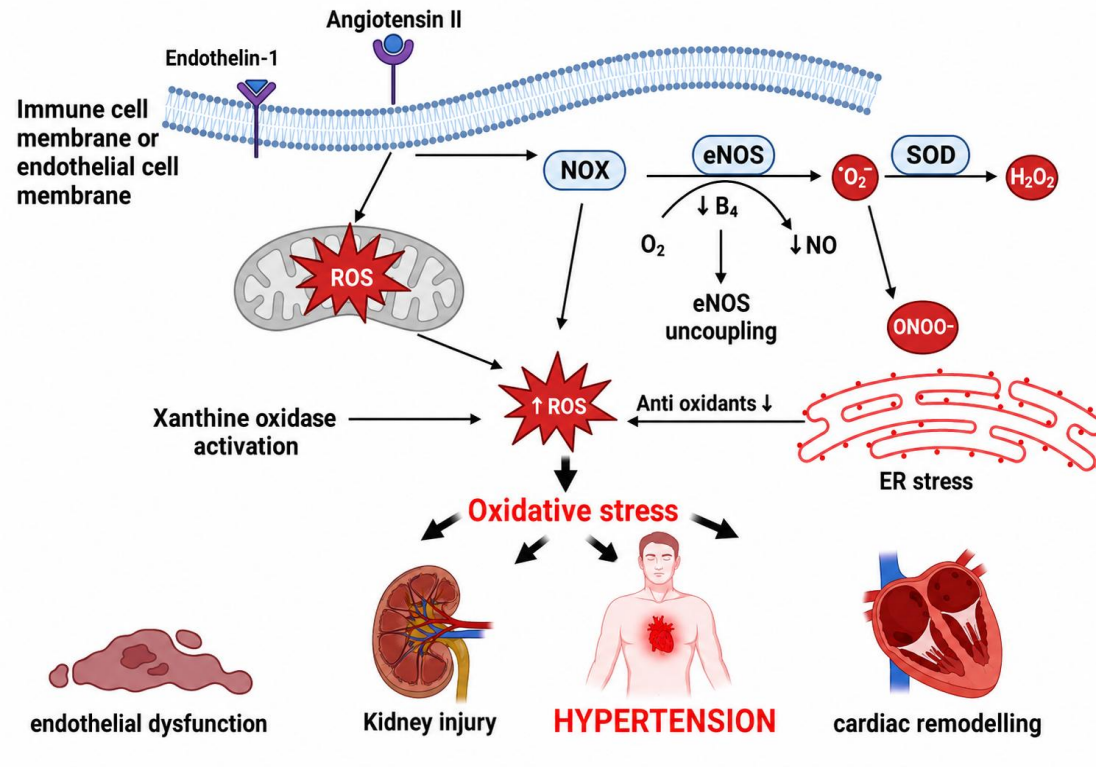


FIGURE 2.3: Oxidative stress pathway showing how NADPH oxidase and mitochondria derived ROS overwhelm antioxidant capacity, causing cellular damage and endothelial dysfunction that reinforce hypertension and cardiac remodeling [56]

There are a variety of downstream effects of this oxidative burden that impact almost every aspect of the syndrome. The subsequent oxidative stress causes endothelial dysfunction and the irreversible buildup of oxidation products, which ultimately reinforce the metabolic syndrome phenotype as a whole by contributing to insulin resistance, hypertension, and dyslipidemia [56].

Oxidative stress markers are being investigated as a therapeutic target for antioxidant based therapies as well as a tool for diagnosing and assessing the severity of metabolic syndrome because biomarkers of oxidative damage can be assessed clinically.

2.1.4 Lipid Metabolism Disorders

A typical pattern of disruption in lipid metabolism, commonly referred to as atherogenic dyslipidemia, is the last fundamental patho-physiological pillar of metabolic syndrome. One of the main characteristics of metabolic syndrome is dyslipidemia, which is defined by altered levels of triglycerides and HDL cholesterol. Elevated triglyceride levels are frequently linked to lower HDL cholesterol concentrations, creating a lipid profile that encourages the development of vascular disease [58].

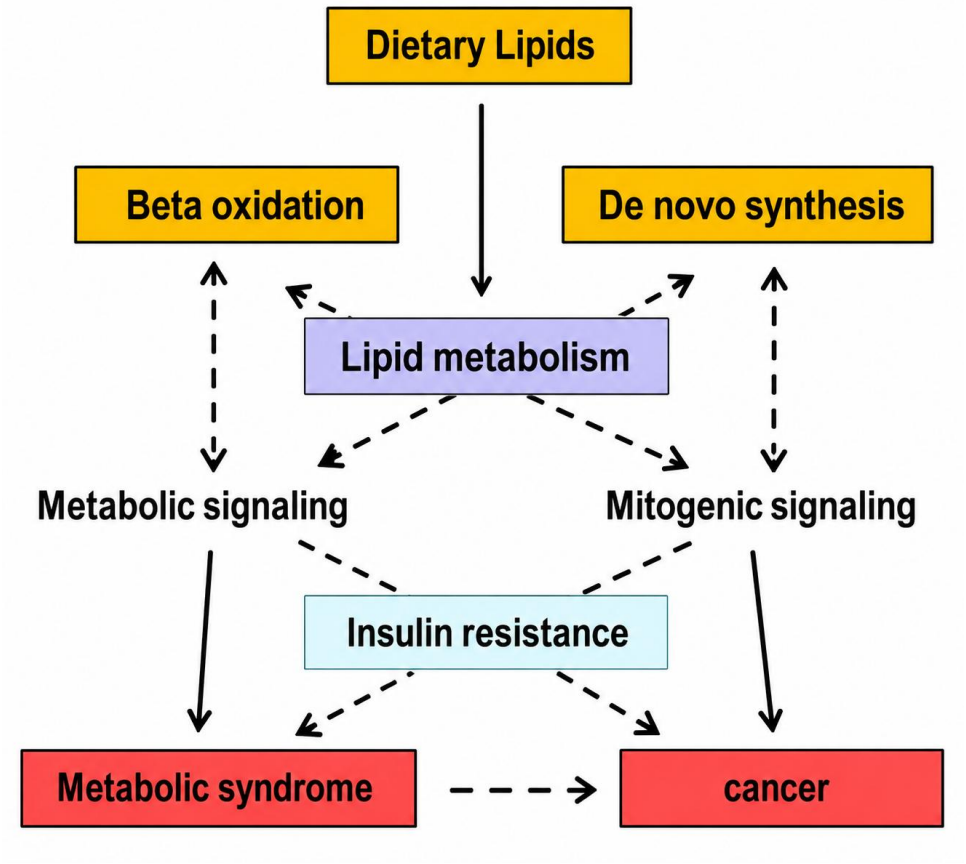


FIGURE 2.4: Interrelationship of Lipid metabolism disorder and Metabolic Syndrome [59].

The previously mentioned insulin resistance is the direct cause of this lipid abnormality. The overproduction of large very low-density lipoprotein (VLDL) particles is caused by impaired insulin signaling, which decreases the normal suppression of lipolysis and increases the flux of free fatty acids to the liver. Together with impaired clearance mechanisms, this fundamental defect is now recognized as the hallmark of the dyslipidemia seen in metabolic syndrome and type 2 diabetes [58].

These VLDL particles mix with HDL and LDL particles throughout circulation, resulting in the production of smaller, denser, and more atherogenic lipoprotein subtypes.

This lipid group has significant clinical implications. A pooled meta-analysis has demonstrated a substantial correlation between novel atherogenic lipid indices, such as non-HDL cholesterol and the triglyceride-to-HDL ratio, and metabolic syndrome and its constituent parts across several diagnostic criteria [60]. This dyslipidemia pattern is considered not only a downstream marker of metabolic syndrome but also an active contributor to the cardiovascular risk that characterizes the syndrome's clinical significance because it directly promotes atherosclerosis and occurs several years before overt diabetes diagnosis.

2.2 Biomarkers of Metabolic Syndrome

2.2.1 Characteristics of an Ideal Biomarker

In contrast to traditional clinical measures like blood pressure, fasting glucose, and lipid profiles, which are helpful but might not fully capture the underlying metabolic disturbances occurring, biomarkers are meant to provide a more objective and accurate assessment of the physiological changes that occur in metabolic syndrome [39]. Therefore, rather than only characterizing the disease's clinical endpoints, an ideal biomarker for metabolic syndrome is anticipated to reflect the biochemical and molecular processes driving the syndrome, notably inflammation, oxidative stress, lipid metabolism, and glucose regulation [39].

A valuable biomarker must be effective as a diagnostic tool in addition to its biological significance. The area under the curve, along with sensitivity, specificity, and an ideal cut-off value, has been used in studies evaluating insulin-resistance-related biomarkers for metabolic syndrome to assess diagnostic potency. Lipid ratios and some non-HDL-cholesterol-based indices have the highest area-under-curve values of 0.80 or above for predicting the syndrome [61]. The assumption

that an ideal biomarker should be repeatable across populations and able to differentiate ill individuals from unaffected ones with a specified, clinically actionable threshold is reflected in this emphasis on quantifiable diagnostic accuracy.

Another distinguishing feature is the practicality of measuring. An ideal marker should be practical to measure on a regular basis, such as through routine blood or urine sampling, and sensitive enough to identify early, subclinical disturbances before the syndrome's clinical components fully manifest. This is because measuring metabolic syndrome biomarkers can improve the diagnostic specificity for each individual and provide clinicians with pertinent clinical information [61]. However, since biochemical marker levels differ among populations and a biomarker validated in one group may not transfer directly to another without recalibration, prediction models based on these markers must be flexible across various ethnic and demographic groups [61].

2.2.2 Clinical Applications of Biomarkers

Metabolic syndrome biomarkers are usually classified into multiple functional categories in clinical and research practice, each of which captures a distinct aspect of the pathophysiology of the syndrome. Insulin resistance markers like insulin, C-peptide, and the HOMA-IR index; inflammatory markers like C-reactive protein, interleukin-6, and tumor necrosis factor-alpha; adipokines like leptin, adiponectin, and chemerin; oxidative stress markers like uric acid and oxidized LDL; vascular markers like arterial pulse wave velocity; and lipoprotein markers like apolipoprotein B [62]. Whether it is glucose dysregulation, systemic inflammation, or atherogenic lipid abnormalities, this classification enables doctors to choose indicators that match to the molecular pathway of greatest interest in a particular patient.

Several of the indicators linked to insulin resistance have demonstrated especially strong clinical correlations. Adiponectin and ghrelin levels were negatively correlated with metabolic syndrome, while fasting insulin, HOMA-IR, leptin, HbA1c, and visfatin levels were positively correlated with the syndrome, according to a systematic review of the Iranian population. The relationship between adiponectin

and the syndrome's individual components was particularly well-established across the literature [61]. The more conventional glucose- and lipid-based diagnostic criteria, adiponectin is increasingly used clinically as a measure of both insulin sensitivity and disease severity because its levels decrease as dysfunction worsens.

Composite indices, which combine many biological signals into a single actionable score, are gaining clinical momentum in addition to individual biomarkers. According to a review of biomarkers of insulin sensitivity and resistance, the triglyceride-to-HDL-cholesterol ratio, the triglyceride-glucose index, and the HOMA-IR index have become useful predictors of insulin resistance, while adipokines like leptin and adiponectin, along with inflammatory mediators like CRP and IL-6, have demonstrated a strong and reliable ability to predict insulin resistance across several prospective cohort studies [63]. Since integrating multiple emerging biomarkers into practice provides a more thorough, customized method of diagnosing metabolic syndrome, tracking disease progression over time, and monitoring a patient's response to treatment, these composite and multi-marker approaches are anticipated to play an increasing clinical role [39].

2.3 DNA Biomarkers

2.3.1 Candidate Genes

A series of genes, with a known biological function in some aspect of glucose, insulin or lipid metabolism, have been found repeatedly and are considered candidates for the metabolic syndrome. These are genes involved in insulin signalling, including insulin receptor substrate 1 (IRS1), peroxisome proliferator-activated receptor gamma (PPARG), as well as genes that are core to lipid metabolism, such as adiponectin (ADIPOQ), apolipoprotein A5 (APOA5) and low-density lipoprotein receptor (LDLR) [64].

Of these, PPARG and APOC3 have been especially consistently associated with metabolic syndrome in multiple independent samples; other candidates like the

genes coding for angiotensin-I-converting enzyme (ACE) and fatty-acid-binding protein 2 (FABP2) have yielded less consistent results [64]. The gene APOA5 in particular has been one of the most replicated candidates with its known role in insulin resistance, systolic blood pressure and triglycerides [65]. These candidate genes were chosen because of prior biological knowledge, and so represent a hypothesis-driven complement to the genome-wide unbiased search for discovery that is used in genome-wide association studies.

2.3.2 Genetic Variants

Based on these "candidate genes", researchers have found specific single nucleotide polymorphisms (SNPs) that have been consistently associated with the risk of metabolic syndrome. A systematic review identified a few SNPs that had minor alleles that were more prevalent in MS patients and minor alleles that were less prevalent in MS patients such as the Taq-1B variant in CETP, multiple variants in APOA5 and APOC3 [64]. Despite some variations in the strength of this association across different ethnic populations and study designs, recent meta-analyses have maintained the consistent association between APOA5 polymorphisms and the susceptibility to metabolic syndrome [66].

Similarly, SNP-based analyses in Middle Eastern cohorts have identified polymorphisms of candidate genes as linked to susceptibility to MS in Middle Eastern populations, although the magnitude and direction of the effects vary across the ethnic groups analyzed [65]. Single SNPs, however, often provide only a small portion of the risk, so nowadays most researchers consider them to be part of a larger genetic risk profile.

2.3.3 Genome-Wide Association Studies

MetaSNP studies have greatly enriched the number of polymorphisms linked to metabolic syndrome beyond the scope of candidate gene studies. In the Taiwan Biobank, a large-scale GWAS identified 549 SNPs, representing 10 genomic risk

loci and 22 associated genes such as CETP, LPL, APOA5, FTO, and GCK, that were significant for metabolic syndrome, with a sample size of over 107,000 individuals [48].

On an international scale, a multivariate GWAS model with a shared genetic factor model for five metabolic syndrome components (fasting glucose, HDL cholesterol, systolic blood pressure, triglycerides, and waist circumference) resulted in the identification of 235 genetic loci associated with metabolic syndrome, compared with 174 genetic loci identified in the largest single-trait GWAS of metabolic syndrome found to date [67].

This expanded discovery is an example of the benefit of modelling the components of the syndrome rather than modelling each risk factor individually, as this approach allows to capture common genetic effects which will not be captured by the single-trait approach. Most of the loci so far found using GWAS methods map to genes previously linked to lipid metabolism and glucose regulation, indicating that these large-scale studies have mostly confirmed, not refuted, the picture of the mechanisms arising from earlier candidate gene studies.

The documented genetic basis of MetS involves a complex array of chromosomal positions and single nucleotide polymorphisms (SNPs) within numerous candidate genes. Positive genetic associations have been successfully replicated for common SNPs in the APOC3 and PPARG genes. Furthermore, specific variants predominantly linked to lipid metabolism demonstrate significant correlations, such as the increased prevalence of minor alleles in FTO (rs9939609), TCF7L2 (rs7903146), APOA5 (C56G, T1131C), APOC3 (C482T, C455T), and IL6 (174G > C) among MetS patients.

Additional evidence implicates the ADIPOQ rs2241766 G-allele with augmented MetS risk in certain populations, alongside sex-dependent associations in the ADRB2 gene for men and correlations with the LMNA 1908C > T polymorphism. Despite these targeted discoveries, the multifactorial heritability and extreme heterogeneity of the syndrome continue to impede the translation of these diverse genetic linkages into routine clinical diagnostic tools [23].

2.4 RNA Biomarkers

2.4.1 Messenger RNA

Messenger RNA expression profiling provides a direct glimpse into the genes currently being transcribed in a patient's cells, and is therefore an obvious next step to consider for biosignature discovery in metabolic syndrome. Liu et al., 2021a identified two gene modules and 51 hub genes associated with metabolic syndrome using peripheral blood high-throughput transcriptomics, and a signature of nine hub genes was distilled, which was used to build a logistic-regression-based risk calculator.

Such an mRNA nine-gene signature showed excellent classifying ability in the training set (AUC 0.968) and in an external validation set (AUC 0.861) suggesting that blood-based mRNA panels can classify patients from non-affected patients with a high level of accuracy [68].

Whereas DNA-based variants offer a snapshot of a predisposition, mRNA-based biomarkers are more suitable to reflect the active, ongoing pathophysiological state of the syndrome–complementing a snapshot of fixed genetic predisposition.

2.4.2 MicroRNA

MicroRNAs are small, non-coding RNA molecules that post-transcriptionally modulate the activity of genes, and their differential expression in the blood has led to their being sought as blood markers of metabolic syndrome. Based on high-throughput sequencing of plasma samples, 2 microRNAs were found to be upregulated and 36 to be downregulated in patients with MS, with two microRNAs being upregulated and 10 being most downregulated being confirmed by quantitative PCR as potential biomarkers [69].

In addition to diagnostic classification, microRNAs are also being investigated to determine whether they can differentiate between similar but distinct metabolic

conditions, such as the presence of a potential biomarker in the plasma of patients with obesity who were not detected in patients who developed metabolic syndrome, indicating that they may be useful for monitoring disease progression instead of just presence or absence [70].

MicroRNAs are highly stable in the blood and can be measured non-invasively and therefore have been proposed as good candidates for standard clinical diagnostic panels, provided their signatures can be reproduced in a consistent manner across different populations.

2.4.3 Long Non-coding RNA

A more general and understudied group of regulatory transcripts are long non-coding RNAs, which interact with other regulatory proteins and microRNAs to regulate the expression of other genes.

In a dedicated review, it was established that metabolism disorders, such as diabetes, obesity, and cardiovascular disease are closely associated with dysregulation of specific lncRNAs, with metabolic disorders being exacerbated by such dysregulation and thus lncRNAs are proposed as diagnostic and therapeutic targets for metabolic syndrome [71].

A large part of this regulatory control takes place via what is termed an lncRNA-microRNA-regulatory-protein axis, where lncRNAs serve as molecular scaffolds or sponges that regulate the availability of microRNAs for their downstream gene targets, positioning lncRNAs upstream of the microRNA biomarkers described above [71].

Although the field has not as yet progressed as far as mRNA or microRNA biomarkers in our understanding of how exactly lncRNAs function in metabolic syndrome, they are considered to be biologically more specific, due to their tissue-specific and context-specific expression.

2.4.4 Circular RNA

Circular RNAs are a novel and unique RNA structure generated by back-splicing that lacks free ends in linear transcripts. CircRNAs are resistant to degradation by exonucleases, which make them more stable than linear RNA species, and this property has led to an increased interest in the potential of circRNAs as circulating, stable biomarkers of metabolic diseases [24].

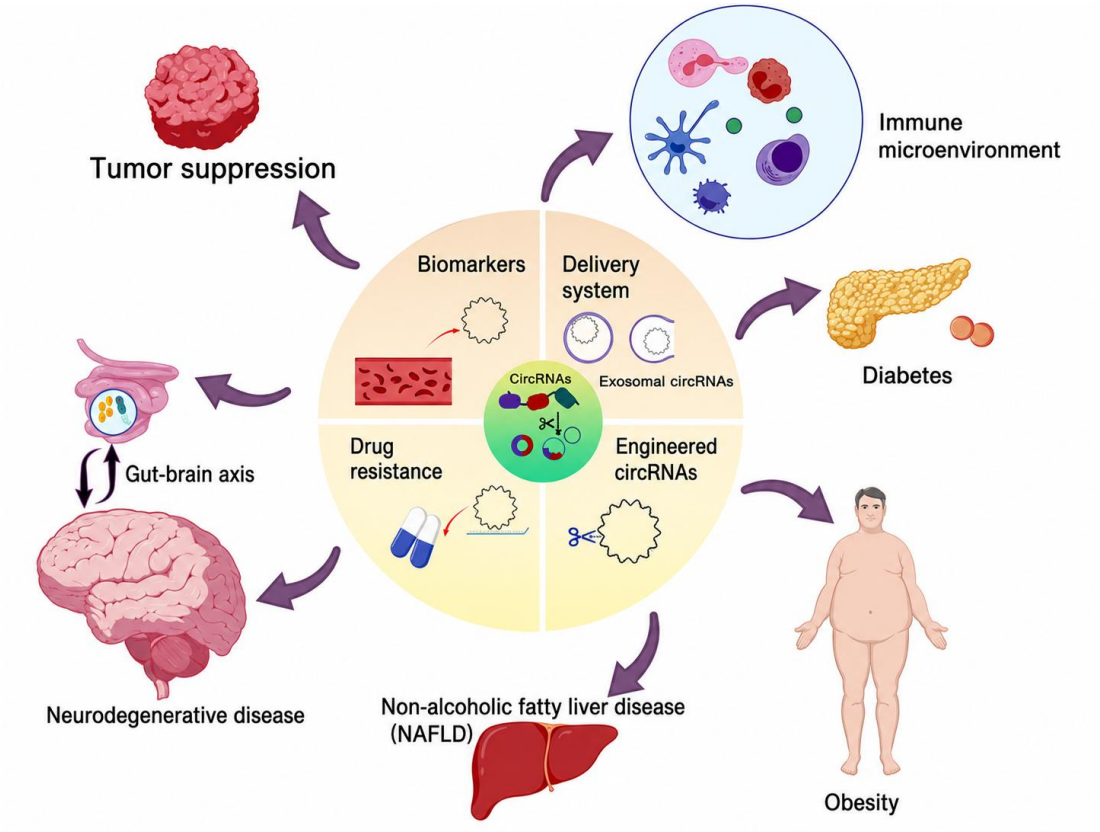


FIGURE 2.5: RNA biomarker in metabolic syndrome [72].

Functionally, circRNAs are known to regulate metabolism basically in two ways: as microRNA sponges that can competitively bind microRNAs via microRNA recognition elements and thus modulate the post-transcriptional regulation of the microRNA target genes [70]; as protein scaffolds that bind proteins and, in some cases, serve as templates for peptide translation [70]. The closed-loop structure makes circRNAs an attractive target for eventual clinical assay development and in recent years they've become one of the most promising of the RNA biomarker classes, though, like the other non-coding RNA biomarkers, correlation

with metabolic syndrome does not demonstrate a causal role of these molecules in the pathogenesis of this syndrome.

2.5 Protein Biomarkers

2.5.1 Inflammatory Proteins

Circulating protein biomarkers are some of the more commonly measured proteins in metabolic syndrome research as the syndrome is tightly linked to chronic low-grade inflammation. A cross-sectional study in middle aged women showed that level of IL-6 was highly correlated with hypertriglyceridemia and insulin resistance, while the level of tumor necrosis factor-alpha (TNFα) correlated with the abdominal obesity and was a significant predictor of cognitive impairment in middle aged women [73].

Interestingly, the same study revealed that both IL-6 and TNF-alpha were linked to specific aspects of the syndrome and did not interchangeably represent the same syndrome; that is, IL-6 was more strongly associated with lipid abnormalities and insulin resistance, and TNF-alpha was more strongly associated with adiposity and its downstream cognitive outcomes [73]. The distinction is clinically relevant as it implies that a single inflammatory marker is not sufficient to account for the inflammatory burden of the syndrome and combinations of more than one cytokine can better reflect the specific metabolic disturbances which occur in a particular patient.

2.5.2 Hormonal Biomarkers

Hormonal biomarkers reflect the neuroendocrine responses and stress-related pathways that are associated with metabolic syndrome, and are not limited to those hormones that are more well known and discussed in relation to insulin. Systematic review of mediators of allostatic load revealed that primary mediators of the

hypothalamic-pituitary-adrenal axis and sympathetic nervous system, such as cortisol, dehydroepiandrosterone sulfate, and norepinephrine and epinephrine, play a role in the development of metabolic syndrome via their interaction with the neuroendocrine system [74].

Another relevant category of hormone biomarkers is the thyroid hormone status, as population-based studies have suggested a sex-specific link between subclinical hypothyroidism and incident metabolic syndrome, and a larger meta-analysis has confirmed that metabolic syndrome was strongly associated with the incidence of thyroid disease overall [75].

Hormonal biomarkers in this category are recognized as important regulators of energy metabolism, lipid handling and blood pressure regulation, and are now increasingly considered as markers of upstream regulatory disturbances that could be responsible for – or part of the explanation for – the more commonly measured metabolic aspects of the syndrome.

2.5.3 Enzymatic Biomarkers

Enzymatic biomarkers such as alanine aminotransferase (ALT), aspartate aminotransferase (AST), and gamma-glutamyl transferase (GGT) have become known as easily available and clinically relevant metabolic syndrome markers. A systematic review and meta-analysis was conducted comparing liver enzyme levels between more than 76,000 patients with metabolic syndrome and over 200,000 without metabolic syndrome, which showed the close relation between liver enzyme levels and metabolic syndrome, and its association was found [76].

These enzymes are part of established liver function tests and their elevation, even in the normal reference range, provides a readily available and cheap marker that could help identify those more likely to have metabolic syndrome before other metabolic markers become elevated. In a clinical environment where more specific inflammatory or genetic biomarker panels may not be readily available, this diagnostic value is thought to be particularly useful.

2.5.4 Adipokines

Adipokines are hormone-like proteins directly released by adipose tissue, and their levels change in a characteristic manner when adipose tissue becomes dysregulated in MS. Leptin, resistin and chemerin are pro-inflammatory adipokines linked to obesity, while adiponectin is anti-inflammatory and decreases in obesity, reflecting the overall shift of adipose tissue towards active inflammatory endocrine organ [53]. The individual adipokines studied have shown the most consistent and well-established relationship with metabolic syndrome, with lower levels of circulating adipokines being well correlated with the individual components of metabolic syndrome in multiple studies, and particularly in the case of adiponectin [61].

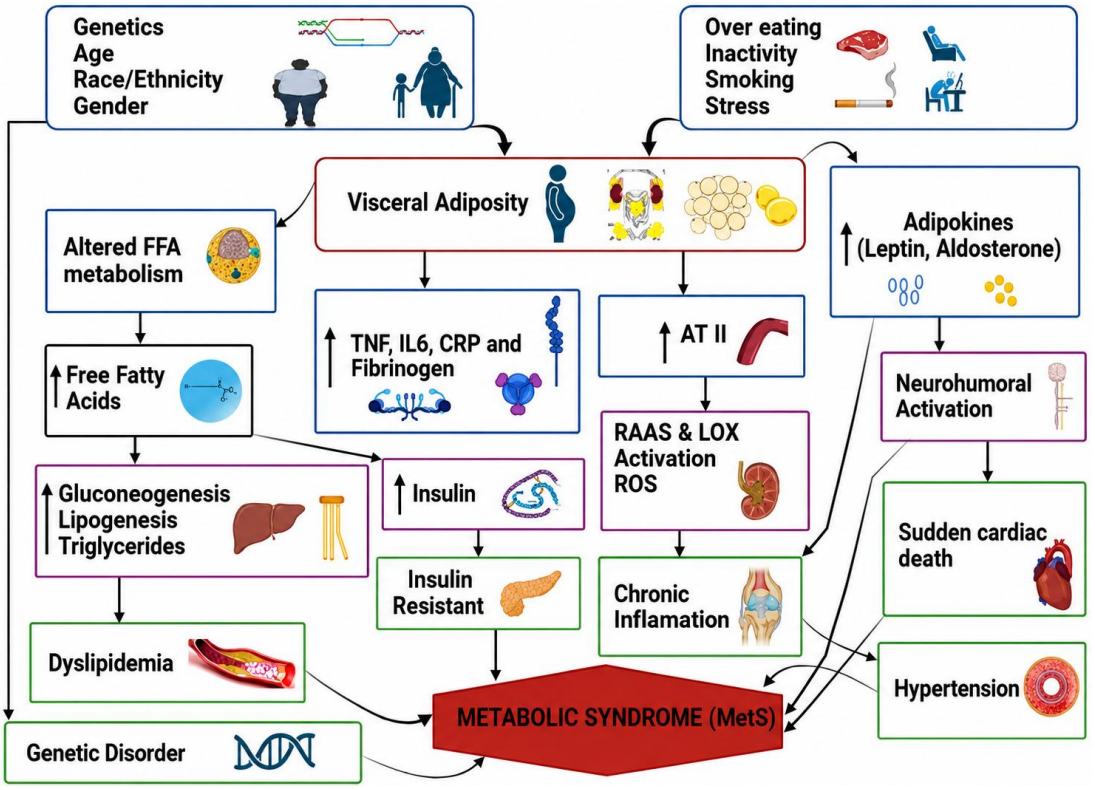


FIGURE 2.6: Schematic presentation of MetS [78].

The levels of adipokines change in predictable ways as visceral adiposity increases, which is why this class of marker is considered to be one of the more mechanistically direct markers to reflect a patient's degree of adipose tissue dysfunction, alongside the more downstream inflammatory and enzymatic markers described above.Metabolic syndrome (MetS) refers to a cluster of clinical features such as central obesity, insulin resistance, dyslipidemia, and hypertension, which together

act synergistically to increase the risk of cardiovascular disease (CVD) and type 2 diabetes. Central obesity involves the accumulation of excess fat in the abdominal region. Insulin resistance represents a key underlying feature and is typically identified through elevated fasting blood glucose levels or abnormal results on glucose tolerance testing [40]. Dyslipidemia in this context is defined by distinct lipid abnormalities, including elevated triglycerides, increased low-density lipoprotein cholesterol, and reduced high-density lipoprotein cholesterol. These alterations contribute to pro-inflammatory and pro-thrombotic conditions, which may be reflected clinically through raised serum biomarkers such as C-reactive protein and fibrinogen. At present, diagnostic definitions differ across professional organizations, although they commonly depend on threshold values for waist circumference, blood pressure, fasting plasma glucose, triglycerides, and HDL cholesterol is a [77].

The pathophysiology of metabolic syndrome (MetS) is primarily driven by dysfunction in adipose tissue, alongside processes such as atherosclerosis, oxidative stress, and persistent low-grade inflammation. Visceral adipose tissue functions as an active endocrine organ, releasing pro-inflammatory cytokines including tumor necrosis factor-alpha (TNF-α), interleukin-6 (IL-6), and interleukin-1 beta (IL-1β) [23]. Furthermore, visceral adipose tissue dysregulation severely alters key adipokine biomarkers; it induces leptin resistance while suppressing the production of adiponectin, an insulin-sensitizing and anti-inflammatory hormone. This ongoing inflammatory milieu induces endothelial dysfunction and stimulates the liver to release acute-phase proteins like CRP, serving as both a clinical marker of inflammation and a contributor to plaque instability [79].

2.6 Epigenetic Biomarkers

2.6.1 DNA Methylation

The most well studied epigenetic modification in metabolic syndrome is DNA methylation, partly because arrays for genome-wide methylation can be readily

used to measure its changes. A study was performed on the Infinium Methylation EPIC BeadChip to look at the differentially methylated probes between people with metabolic syndrome and controls and found that these probes predominantly clustered in gene bodies and intergenic regions, while hypomethylated probes tended to cluster near the transcription start sites [80].

In the genes identified, GFPT2 hypermethylation was a very interesting candidate as a possible biomarker, and the authors of the study state that the aberrant expression of GFPT2 could be a direct mechanism of the syndrome's pathogenesis, rather than an indirect correlation [80]. DNA methylation patterns appear to be good candidates for clinical diagnostic assays as they can be extracted from accessible tissues like peripheral blood, and are relatively stable over short periods of time; additional validation cohorts are required to determine which sites reflect methylation across the population.

2.6.2 Histone Modifications

The second major epigenetic mechanism is histone modification, in which the DNA is not modified, but rather the tightness of its packaging around histones is altered to regulate gene expression. The authors of a review on adipogenesis and metabolic syndrome highlighted the importance of histone modification, in addition to DNA methylation and chromatin remodeling, in adipocyte differentiation and how these epigenetic changes are influenced by environmental factors including energy-dense diets, altered sleep patterns and sedentary lifestyle [81].

Histone marks are considered to be a mechanistic link between lifestyle related risk factors and the overall gene expression changes that culminate in the metabolic syndrome phenotype, as they depend on the interaction between an individual's genes and other environmental factors [81]. The environmental sensitivity also implies that, in principle, histone modifications are reversible, which has led to speculation of their use as biomarkers of accumulated metabolic risk and targets for therapeutic intervention.

2.6.3 Epigenetic Regulation in Metabolic Syndrome

All together, DNA methylation and histone modifications represent part of a larger epigenetic regulation system that includes chromatin remodeling and non-coding RNA mechanisms, targeting the same metabolic pathways downstream. In a detailed review of epigenetics in metabolic disease, these four mechanisms were identified as the main mechanisms by which epigenetics are modulated, and it was found that these epigenetic layers are modified by genetic and non-genetic factors together to create an individual's metabolic phenotype [30]. The integrative perspective has immediate clinical relevance, as already demonstrated in the clinic through epigenetic biomarkers, drugs, and new epigenetic editing technologies [30]. Epigenetic regulation, which lies on a mechanistically intermediate step between genetic risk and the molecular changes at transcript level and protein level outlined above, is therefore becoming increasingly recognized as a layer that links a person's genetic risk with the measurable molecular profile of metabolic syndrome.

2.7 Complications

The synergistic effect of MetS components significantly increases the odds of developing severe cardiovascular diseases, including coronary artery disease, stroke, and peripheral artery disease. Type 2 diabetes mellitus (T2DM) is a primary complication resulting directly from the core feature of insulin resistance, which is further exacerbated by inflammatory cytokines and free fatty acids released from excess adipose tissue [82]. Metabolic syndrome (MetS) is strongly associated with metabolic dysfunction-associated steatotic liver disease (MASLD), which develops from overlapping abnormalities including central obesity, insulin resistance, and dyslipidemia. It is also linked to additional comorbidities such as polycystic ovary syndrome (PCOS) in females, in which hyperinsulinemia promotes excessive ovarian androgen synthesis, as well as chronic kidney disease, where persistent hypertension and hyperglycemia contribute to glomerular hyperfiltration and progressive renal damage [19].

2.8 Management and Therapeutic Approaches

Effective prevention and treatment of MetS necessitate addressing its multitudinous risk factors simultaneously through lifestyle modifications and targeted pharmacological interventions. Physical activity—encompassing aerobic exercise, high-intensity interval training, and strength training—mitigates MetS by improving insulin sensitivity, stimulating mitochondrial biogenesis, and enhancing fat metabolism. Dietary therapy, such as the Mediterranean diet, plays a central role by utilizing high-fiber foods, essential fatty acids, and low refined carbohydrates to stabilize blood sugar, reduce oxidative stress, and lower inflammatory cytokines. When lifestyle modifications are insufficient, specific medications are utilized to target individual MetS biomarkers, including Metformin and Thiazolidinediones for glycemic control, statins and fibrates for managing LDL and triglyceride levels, and antihypertensive drugs for blood pressure regulation [48].

2.9 Biomarker Discovery Technologies

2.9.1 High-Throughput Sequencing

The traditional candidate-gene approach is giving way to the new high-throughput sequencing technology for unbiased and genome-wide discovery of biomarkers in metabolic syndrome. Such an approach was applied to plasma samples, resulting in the identification of differentially expressed microRNAs in healthy volunteers as compared to those with metabolic syndrome or MS, which were then further reduced by bioinformatic enrichment analysis and experimentally validated using quantitative PCR [69].

The workflow, from wide sequencing to targeted validation, has now also been adopted as the methodological paradigm for RNA and DNA-based biomarker discovery in metabolic syndrome, as it permits to cast a large net and then target resources on those that seem the most promising.

2.9.2 Proteomics

Proteomic technologies go beyond DNA- or RNA-based technologies to discovering biomarkers at the protein level, adding another layer of biological information to the genome and transcriptome that cannot be fully predicted. Along with genetic and transcriptomic and metabolomic analyses, a review of multi-omics profiling in metabolic disease highlighted proteomics as one of the tools to be used for more accurate characterization of the pathogenesis of metabolic syndrome and related metabolic disorders, aiming to move towards more precise, individual treatment strategies [83]. Proteomic data do not simply replicate transcriptomic data, but rather provide a more direct readout of the functional molecules, enzymes, cytokines and hormones that ultimately give rise to the clinical manifestations of the syndrome and should therefore be regarded as a valuable complement to transcriptomic data.

2.9.3 Transcriptomics

A more comprehensive analysis of the totality of RNA transcripts in a sample can be achieved using transcriptomic technologies, and this can give a genome-wide picture of the activity of genes that underlies the more targeted mRNA biomarker panels mentioned above. Based on the largest peripheral blood transcriptomics data available to date, which encompasses 137,962 samples from peripheral blood for individuals with metabolic syndrome, researchers used a weighted gene co-expression network analysis in combination with protein-protein interaction network analysis, LASSO regression and random forest methods to narrow down a list of 51 possible hub genes to a nine-gene signature [68].

This multi-layered analytical approach, which integrates multiple independent statistical and machine-learning techniques to arrive at a small, robust set of genes, represents one of the ways in which transcriptomic technologies have progressed beyond the mere identification and reporting of genes that are differentially expressed to the development of a truly predictive, clinically actionable gene panel.

2.9.4 Epigenomics

Epigenomic technologies add to the omics toolkit by profiling the patterns of methylation, modification of histones, and chromatin accessibility across the genome in a large-scale, high throughput fashion, placing epigenetic biomarker discovery in the same context as genomic and transcriptomic profiling. The authors of the editorial, "Integrated multi-omic studies of metabolic syndrome and related insulin disorders," identified epigenomics, a study of gene expression regulation, as one of the key molecular layers in addition to genomics, transcriptomics, proteomics, microbiomics and metabolomics, which together will uncover the complex mechanisms and interactions that underlie metabolic disease [84].

The multi-omics approach, which merges the genetic, transcriptomic, and proteomic technologies outlined in the previous subsections with the epigenomic approach, is becoming the most promising strategy to develop the type of precise and personalized biomarker panels required for early detection and targeted intervention for the syndrome.

2.10 Existing Biomarker Databases

2.10.1 National Center for Biotechnology Information

One of the most frequently used collections of online resources for biomarker research is provided by the National Center for Biotechnology Information (NCBI), providing access to biological data through search and retrieval via 35 distinct interconnected databases including GenBank and PubMed, via the shared E-utilities programming interface [85]. In addition to storing raw sequence and literature data, NCBI resources like ClinVar, dbGaP, and the Genetic Testing Registry (GTR) include clinically oriented layers of curation that enable straightforward linking of a genetic variant to the disease phenotype, so that NCBI offers a logical starting point for research screening candidate metabolic syndrome biomarkers for existing genomic and clinical evidence [85].

2.10.2 Ensembl

For over 20 years, Ensembl has created high-quality genome resources for vertebrates and model species, providing a platform for researchers to access annotated reference genomes, comparative genomics tools and a widely-used Variant Effect Predictor (VEP) for interpreting the functional consequences of genetic variants [86]. In the context of metabolic syndrome biomarker discovery, Ensembl's ability to integrate data from the population-scale variation and interactive genome browsers enables researchers to position candidate SNPs and genes including those from GWAS studies earmarked in earlier sections in the entirety of their genomic and regulatory context.

2.10.3 UniProt

The main goal of UniProt is to provide a high-quality, comprehensive and open repository of protein sequences and functional information for the tree of life, now counting over 220 million sequences [87]. In addition to static sequence information, UniProt incorporates large-scale data from clinically relevant variation sources and mass-spectrometry based proteomic evidence, which is directly linked to protein sequence records, thus making it directly relevant to supporting the validation of the protein biomarkers, including adipokines and inflammatory cytokines, discussed in earlier parts of this review [87].

2.10.4 Gene Expression Omnibus

Since its launch over 20 years ago, the Gene Expression Omnibus (GEO) has been an international public repository where gene expression and epigenomics datasets acquired via next-generation sequence and microarray technologies have been deposited and archived, now housing data and links to tens of thousands of published manuscripts [88]. GEO stores raw and processed transcriptomic and epigenomic data submitted directly by the research community, making it the most directly relevant resource for transcriptomic and epigenomic biomarker discovery

methods described above, in order to reanalyze existing sets of data on expression of metabolic syndrome, or to submit raw and processed data from their own studies for validation by others.

2.10.5 Online Mendelian Inheritance in Man

Online Mendelian Inheritance in Man (OMIM) is a large, regularly updated database of human genes and genetic diseases, dedicated to the molecular connection between variation and expression, and only entries curated from and cited to the peer-reviewed biomedical literature [89]. In the context of metabolic syndrome research, OMIM is a tool to connect candidate genes mentioned in previous sections to the known gene-phenotype relationship, serving as a curated link between a specific genetic variant and its known clinical relevance.

2.10.6 Other Disease-Specific Biomarker Databases

Aside from these widely used data-bases, a few others are focused on metabolic biomarkers or metabolic syndrome. The Human Metabolome Database (HMDB) offers detailed information about small-molecule metabolites occurring in the human body, and is specifically designed for metabolomics, clinical chemistry, and discovery of metabolite biomarkers; it has been used directly to identify candidate metabolite biomarkers of metabolic syndrome and its five components in population-based cohorts [44]. More directly, the Metabolic Syndrome Research Resource (MetSRR) is the curation of current and potential biomarkers extracted from the National Health and Nutrition Examination Survey (NHANES) database of $>$ 90,000 individuals to explore the etiology of MetS across demographic groups, according to the developers, who state that MetSRR is the only MetS-specific database designed to do this [90]. The purpose-built databases like HMDB and MetSRR complement general-purpose ones like NCBI, Ensembl, UniProt, GEO and OMIM, which are not necessarily built around any particular disease, and enables researchers to search directly for biomarkers linked to metabolic syndrome, rather than build that link from more general genomic or proteomic records.

2.10.7 Need for the Proposed Study

The problem statement's limitations: fragmented integration, poor accessibility, and inconsistent standardization, show that the current literature on metabolic syndrome biomarkers is insufficient to support the kind of systematic, cross-dimensional research that precision medicine is increasingly demanding. Researchers are forced to manually reconstruct this evidence base study by study in the absence of a curated resource that compiles molecular, clinical, and phenotypic biomarker data in one location. This process is ineffective, challenging to replicate, and prone to missing relevant findings dispersed across various sources and formats.

The current work is directly motivated by this disparity. The study tackles the fragmentation and lack of standardization in the field at its source, before the data ever reaches a researcher attempting cross-study comparison, by methodically identifying, extracting, and curating experimentally validated biomarkers (Objective 1) and classifying them into standardized DNA, RNA, protein, and epigenetic categories (Objective 2). The study goes beyond curation to include the creation of an interactive, web-based platform since curation alone is insufficient if the final data is still as hard to obtain as the original literature.

Each of the three suggested modules is intended to address a particular weakness noted in the problem statement. By substituting systematic, filterable retrieval for dispersed manual literature searches, the Browse module (Objective 3) directly overcomes the accessibility barrier. By converting individual data points into statistical summaries and visual patterns throughout the selected biomarker collection, the Insights module (Objective 4) addresses the lack of a unified, comparative context. By enabling direct comparison of similar biological characteristics and functional correlations across biomarkers, the Analyse module (Objective 5) goes one step further and addresses the lack of integration in the field. This capacity is not possible with fragmented, single-study reporting alone.

When considered collectively, these goals show that the proposed study is corrective rather than merely supplementary to the body of existing literature; it is

especially intended to address the fragmentation, inaccessibility, and lack of standards that currently limit research on metabolic syndrome biomarkers. In order to facilitate repeatable research, support extensive meta-analysis, and eventually further the translation of biomarker discovery into standardized, precision-medicine-oriented clinical practice, it is imperative that such a platform be established.

2.11 Research Gap

2.11.1 Limitations of Existing Databases

Although there have been great advances in bioinformatics resources, the current biological databases suffer from certain limitations when used for metabolic syndrome (MetS) biomarker research. The majority of the public databases, such as Human Metabolome Database (HMDB), Gene Expression Omnibus (GEO), DisGeNET, STRING and KEGG, are broad-based and contain a wide variety of diseases, not just metabolic syndrome. This makes it difficult to retrieve comprehensive information and integrate it into a single pipeline within a single database, as information about MetS biomarkers is typically spread across many databases. Wishart et al. (2022) pointed out that the HMDB is essentially concerned with the identification of metabolites and metabolic pathways, but lacks disease-specific curation of biomarkers.

Likewise Barrett et al. (2013) explained that GEO is a repository for datasets from functional genomics, which need a lot of preprocessing and bioinformatics expertise before any meaningful biological interpretation can be drawn from the data. Moreover, many databases are not annotated with a standard, do not have stable nomenclature and do not have integrated clinical metadata, and this reduces the replicability and comparability of biomarker studies. While DisGeNET combines from various sources information on gene–disease associations, differences in data quality, evidence types and annotation standards are important issues as observed by [91].

Moreover, the information on the validation status of biomarkers, their diagnostic performance, the type of samples, the study population, and the clinical applicability is rarely available, which hinders researchers and clinicians from selecting reliable biomarkers for metabolic syndrome. These restrictions highlight the importance of more specialized, curated and disease-specific databases which can combine multi-omics data with clinical evidence to facilitate the discovery of biomarkers and translational research.

2.11.2 Need for a Dedicated MetS Biomarker Platform

As it becomes more common and more complex, there is a need for a specialized biomarker database that brings together the validated molecular and clinical information in one convenient, easy-to-use platform. MetS is a multi-faceted metabolic disorder, and diagnosis and prognosis depends on many biomarkers derived from genomics, transcriptomics, proteomics, metabolomics, and clinical studies. At the present time, researchers are required to search several different independent databases and manually perform data integration from heterogeneous sources which is time consuming and error-prone. Hasin et al. (2017) noted that multi-omics data integration is vital for understanding and finding clinically relevant biomarkers of complex diseases, while Subramanian et al. (2020) pointed out that centralized biological databases enhance data access, repeatability and collaborative research.

A specific MetS biomarker database will comprehensively archive experimentally-validated biomarkers and their corresponding potential biological functions, molecular pathways involved, associated genes, protein interactions, diagnostic and prognostic importance, experimental validation status, study population, sample source, and literature references. This would make biomarker prioritization easier and would allow for systems biology analyses, machine learning applications for disease prediction and quick translation of research results into clinical practice. Finally, a highly specialized metabolic syndrome biomarker database would be an important tool for researchers, clinicians and pharmaceutical scientists to

implement precision medicine, provide greater diagnostic accuracy, and develop targeted therapeutic interventions.

2.12 Summary of Literature Review

In summary, the current literature highlights the growing importance of diagnostic biomarkers in understanding and managing metabolic syndrome, while simultaneously revealing significant gaps in data integration, standardization, and accessibility. These insights emphasize the necessity for a comprehensive platform that consolidates and curates biomarker knowledge in a systematic manner. The MetS-Marker framework is designed to bridge these gaps by combining rigorous literature mining with advanced data integration and exploration capabilities. Building upon the limitations identified, the subsequent methodology outlines the processes employed for data collection, curation, and platform development, providing a structured foundation for the implementation of this resource.

Chapter 3

Methodology

3.1 Introduction

The construction workflow of the MetSMarker consists of two major components: literature-based biomarker mining and the development of a biomarker knowledge exploration platform. The first component includes the systematic literature survey, data selection, information extraction, preprocessing, and curation respectively. The second component integrates the ontology annotations, precision medicine knowledge exploration, as well as the multi-dimensional data interpretation. The overall framework ensures comprehensive identification, validation, and utilization of diagnostic biomarkers associated with the metabolic syndrome.

3.2 Literature Survey and Selection Criteria

A wide-ranging literature search was carried out for published reports that reported diagnostic markers for metabolic syndrome. PubMed was used as a major source of biomedical publications for the literature search. A structured search strategy was developed with the aim of gaining maximum searching efficiency by using the terms related to biomarker expressions and the terms associated with metabolic syndrome and diagnostic terms.

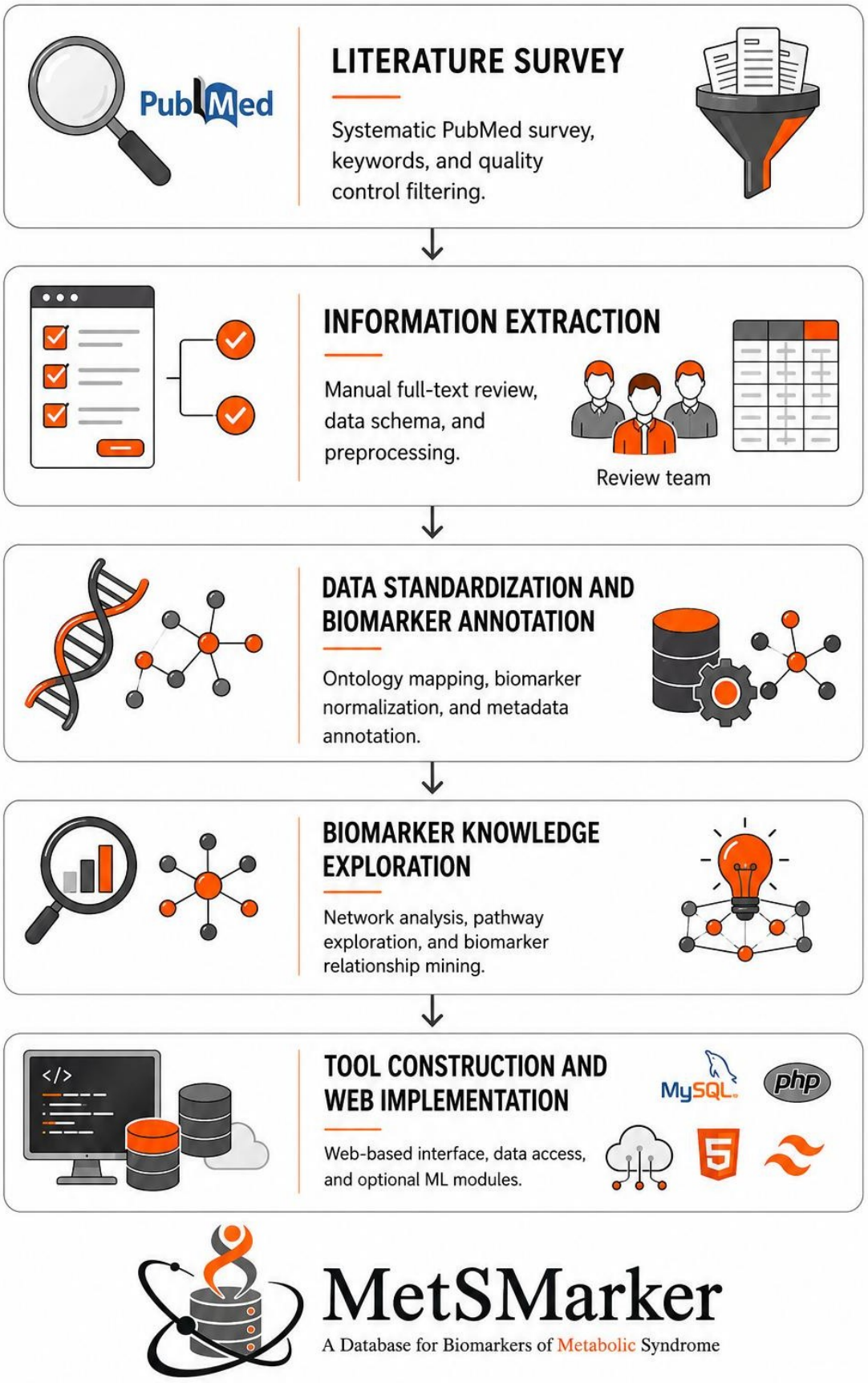


FIGURE 3.1: Methodology flowchart for MetSMaker including literature mining, biomarker curation, ontology annotation, knowledge integration, and multidimensional interpretation.

The following search query was used: ((Biomarker [Title/Abstract]) OR (marker [Title/Abstract]) OR (indicator [Title/Abstract]) OR (predictor[Title/Abstract])) AND ((Metabolic Syndrome [Title/Abstract]) OR (insulin resistance [Title/Abstract]) OR (obesity [Title/Abstract]) OR (dyslipidemia [Title/Abstract]) OR (hypertension [Title/Abstract])) AND ((Diagnosis [Title/Abstract]) OR (diagnostic [Title/Abstract]) OR (Examinations [Title/Abstract])). The first search produced 9,286 publications from 1965 to 2026. The search results were then narrowed to studies that were conducted on humans, resulting in 6,936 publications.

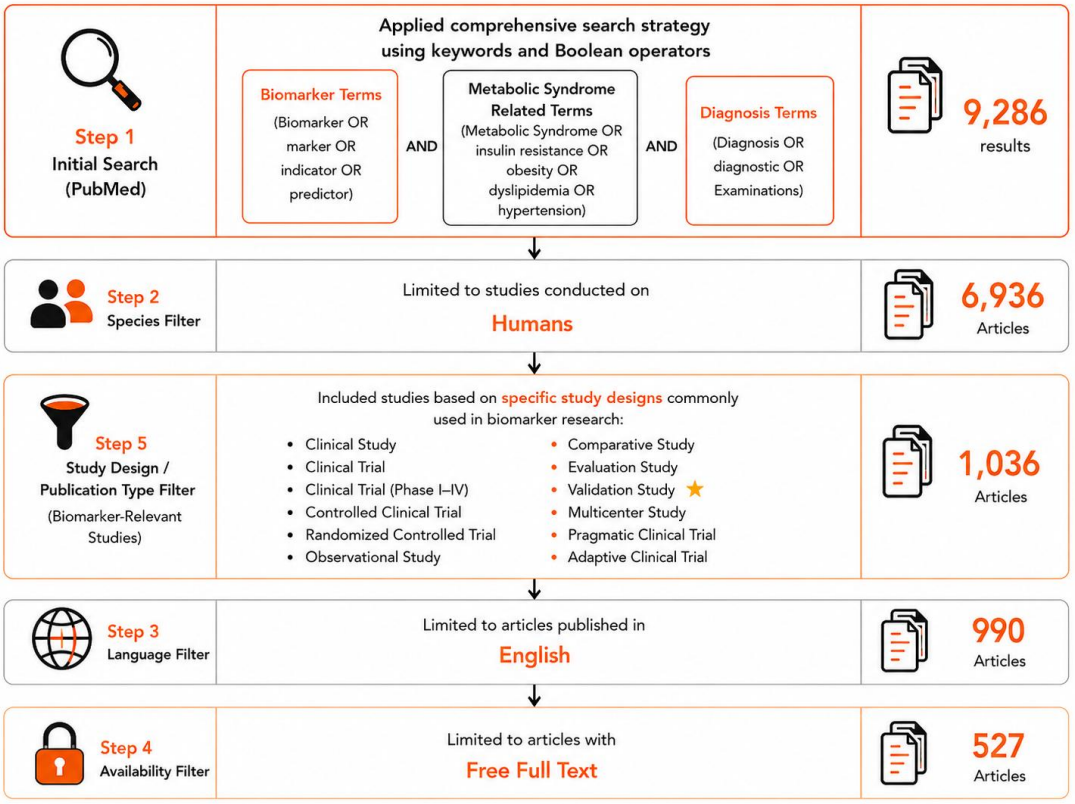


FIGURE 3.2: Literature selection workflow illustrating the sequential filtering strategy used to identify eligible studies, including search retrieval, human subject restriction, publication type selection, language filtering, and free full-text availability.

Further filtering was then done to keep only original clinical investigations. Eligible publications were clinical studies, clinical trials (Phase I through Phase IV), controlled clinical trials, randomized controlled trials, observational studies, comparative studies, evaluation studies, validation studies, multicenter studies, pragmatic clinical trials and adaptive clinical trials. Articles that did not contain any

primary experimental evidence, such as review articles, systematic reviews, meta-analyses, editorials, letters, comments, conference proceedings, book chapters, case reports, clinical guidelines, biographies, interviews, news articles, errata, duplicate publications, retracted publications, and other non-original research articles were excluded. After this filtering step, there were 1,036 articles left that were original research articles.

For uniformity in data extraction and interpretation, only studies published in English language were included, leaving 990 eligible articles. Lastly, all search papers were limited to full-text articles freely available online, leaving 527 studies to be reviewed in detail by hand and for potential extraction of biomarkers.

3.3 Information Extraction

Selected articles were manually curated by independent data reviewers to guarantee data accuracy and reproducibility by reviewing the full text of the article. Information extracted was biomarker characteristics, study design, demographics (age, sex, BMI, blood pressure), assay methods and statistical significance. A structured data schema was created to standardize data throughout studies. Data preprocessing involved removing duplicate entries, normalizing the nomenclature of the biomarkers with the appropriate biomedical ontologies and harmonizing the measurement units. The missing or inconsistent data were addressed using pre-established criteria so as to ensure integrity of the dataset. All biomarkers were listed individually, with multiple entries for the same biomarker when it was validated in various studies. Reliability for the curated data set was checked at the end.

3.3.1 Stranded for extraction

A standard list of data was created for the extraction of data from selected research articles, to make sure it is systematic to collect it and uniform throughout the review area. The following standardized extraction sheet was used as a

template for recording and structuring the information gained from each of the selected studies. The predefined extraction format reduced the variations in data collection and supported systematic comparisons of the results obtained from various publications. The key pieces of data that were found in each of the papers selected were grouped into the following categories:

i. Literature Information (doi, journal, publication_time, title, abstract, keywords): This category consists of bibliographic and publication data for each selected article, so that the literature in the study can be identified, documented and referenced accurately.

ii. Biomarker Information: (gene_symbol, biomarker_description, type_of_biomarker, type_of_rna _biomarker, clinical_use, level_of_evidence) contains information related to the investigated biomarker, its biological description, type, type of rna-biomarker, clinical use and level of evidence reported in the study.

iii. Clinical Information: (research_region, total_number, male, female, mean_age, age, stage) This is a summary category of the demographic and clinical characteristics of the study population that gives the necessary information about the distribution of patients, age, sex, stage of disease, and geographical origin of the research.

iv. Experiment information: (source, key_experiment) This category describes the experimental source and the major experimental strategies used in each experiment or study to explore the reported biomarkers.

v. Knowledge: (up_regulator, down_effector_or_targets, knowledge_points) This category reports any known biological mechanisms associated with each biomarker, such as up or down regulators, downstream effectors or targets, and other important biological information found in the selected articles.

vi. Disease: (disease_type, disease_subtype, disease_classification) This category refers to disease characteristics investigated in each study; disease type, subtype, and classification used by the authors.

vii. Statistics: (sensitivity, specificity, area_under_the_curve, supplementary_statistics) This category covers all the statistical measurements reported in the studies to assess the diagnostic performance of biomarkers, as well as all other supplementary statistical analyses reported in the studies.

For standardization and reproducibility of information extraction, predefined parameters were set for each category and subcategory incorporated into the extraction sheet. These parameters gave a uniform structure to the recording of the same kind of data from all the selected studies and gave better reliability and comparability to the extracted data set. The following set of parameters were set for each field:

i. field_name

ii. category

iii. Data Type

iv. Description

v. Example

These parameters ensured that all the data fields were uniformly defined and extracted, which would make it possible to organize and analyze data in a systematic way and across all the selected articles for uniformity.

3.4 Data Standardization and Biomarker Annotation

After data extraction the curated information was systematically standardized and annotated to make it consistent, accurate and interoperable in the MetS-Marker database. Data standardization consisted of integrating the biomarker-related terms, disease terminology and molecular classification using internationally accepted and recognized standards, and integrating other attributes extracted.

Naming of biomarkers, disease description and reporting format for the different studies were carefully examined and standardized to ensure consistency of the data structure. To ensure the quality and reliability of the curated dataset, duplicate records, inconsistent terminologies and suggestive entries were manually checked and rectified. After standardization, each biomarker was annotated to a set of well-known international biomedical databases to improve the information extracted with standardized identifiers and biological annotations. Biomarker annotations, such as gene information, standardized names and associated biological details, were obtained as the primary source from the NCBI. To validate the information about the biomarkers and to provide additional annotations, other publicly available biological databases were used, when needed. This annotation process provided the database with increased biological relevance, interoperability and usability, allowing for precise identification of biomarkers, data retrieval and exhaustive downstream analysis via the MetSMarker platform.

3.5 Biomarker Knowledge Exploration

The key knowledge exploration modules developed for MetSMarker were: Browse, Insights, and Analysis. These modules enable users to have ready access to the biomarker information, detailed statistics of the database, and biomarker exploration tools.

i. The Browse function allows the user to systematically browse the metabolic syndrome biomarkers with various search criteria such as biomarker term, disease and type of organ associated with the biomarker and category of the biomarkers. This module offers quick access to curated biomarker records and allows for easy access to relevant information based on user defined search parameters.

ii. Insights: Statistical summaries and interactive visualisation of curated database. It allows users to view the distribution of the biomarkers in each category, explore the distribution of each biomarker in the metabolic syndrome,

and provides a summary of the characteristics and contents of the database in graphical and statistical formats.

iii. Analysis module is able to facilitate comparative and cross-omics analysis of metabolic syndrome biomarkers. This module allows users to compare biomarkers across the various molecular categories and studies, and to present the analysis results using detailed comparison tables to aid interpretation and the identification of biologically relevant patterns and relationships.

3.6 Tool Construction and Web Implementation

The web-based database framework was designed with HTML5 (for front-end structure), CSS3 (for the styling of the interface), JavaScript (ES6) (for the interactive part), PHP 8 (for the back-end processing) and SQL (for the management of the database). This data visualization relied on the HTML5 Canvas API, and ECharts v5.4.3, which was integrated with the v4.9.0 World Map JSON. The platform is composed of 3 main modules. The Browse function allows users to search and retrieve the biomarker information by disease, organ, biomarker category or search terms which allows the users to have efficient access to the contents of the database.

All curated insights or selected insights are visible, accessible and can be downloaded from the Insights module. It also includes access to the prevalence data that is associated with it, and displays the geographical distribution with ECharts v5.4.3 and v4.9.0 World Map JSON framework. The Analysis module enables cross-omics comparisons, provides downloadable comparison tables and allows cross-omics analysis via interactive network diagrams embedded in the bioPilot software using Cytoscape.js to visualize node–edge networks. Other analytical graphs and charts were created with the HTML5 Canvas API.

Chapter 4

Results

4.1 Overview of MetSMarker

MetSMarker was successfully evolved to be a complete knowledge platform on the web for metabolic syndrome related biomarkers. A total of 207 biomarkers were manually curated based on published literature from 1973 to 2026. Biomarker information was systematically standardized and annotated with relevant information from the biomedical and clinical domain such as biomarker type, associated diseases, sample source, level of evidence, organ association and external molecular identifiers to facilitate its efficient exploration and retrieval.

Table 4.1: Top 10 most cited biomarkers based literature reviewed

Rank	Biomarker Name	Supporting Articles
1	C-Reactive Protein	13
2	N-Terminal Pro-B-Type Natriuretic Peptide	8
2	Triglyceride-Glucose Index	8
4	Creatinine	6
5	B-Type Natriuretic Peptide	5
5	Placental Growth Factor	5
7	Interleukin 6	4
7	Triglyceride	4
7	Tumor Necrosis Factor Alpha	4
10	Cholesterol	3

The selected markers were those that were reported in the literature with different frequencies (Table 4.1). The most common biomarkers identified were C-reactive

protein (CRP) (13 publications) and N-terminal proB-type natriuretic peptide (NT-proBNP) (8 publications), and the triglyceride-glucose (TyG) index (8 publications). Creatinine was reported in 6 publications, while B-type natriuretic peptide (BNP) and placental growth factor (PlGF) were reported in 5 publications each, and tumor necrosis factor-alpha (TNF-α) was reported in 4 publications, while the other proteins were reported in less than 5 publications each.

The platform covered a broad range of molecular diversity as shown by the other recurrent biomarkers, such as IL-6, adiponectin, cholesterol, HbA1c, intestinal fatty acid-binding protein (I-FABP), retinol-binding protein 4 (RBP-4), triglycerides, high-sensitivity C-reactive protein (hs-CRP), soluble ST2 (sST2), fatty liver index (FLI), cystatin C, cytokines, and endostatin.

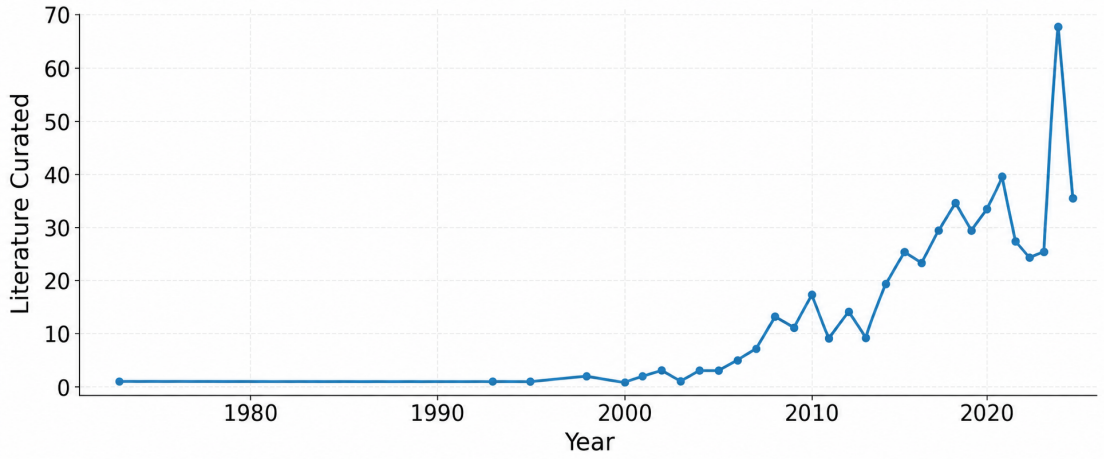


FIGURE 4.1: The analysis of publication trends, which demonstrates a progressive increase in the metabolic syndrome biomarker research over the time.

Trends in publications showed a gradual rise in publications on metabolic syndrome biomarkers over time (Figure 4.1). The highest number of included publications was recorded in 2025 (68 studies), followed by 2021 (39 studies), 2026 (35 studies), 2018 (34 studies), and 2020 (33 studies).

Prior to 2005, relatively few studies were published, but over the past decade, there was a significant increase in the number of studies published for the discovery and validation of biomarkers for metabolic syndrome.

4.2 Characteristics of Curated Biomarkers

4.2.1 Biomarker Types

All 207 biomarkers associated with metabolic syndrome were curated and grouped into six major molecular categories (Figure 4.2). Protein biomarkers were the highest in number (95, 45.9%) and metabolic biomarkers were the second highest (56, 27.1%). Out of the 207 total, 18 (8.7%) were clinical biomarkers and 17 (8.2%) were RNA biomarkers.

Fourteen (6.8%) entries were immunological markers and 7 (3.4%) entries were DNA markers. The distribution shows that the majority of the experimentally validated biomarkers, which are included in MetSMarker, are proteins and metabolic markers.

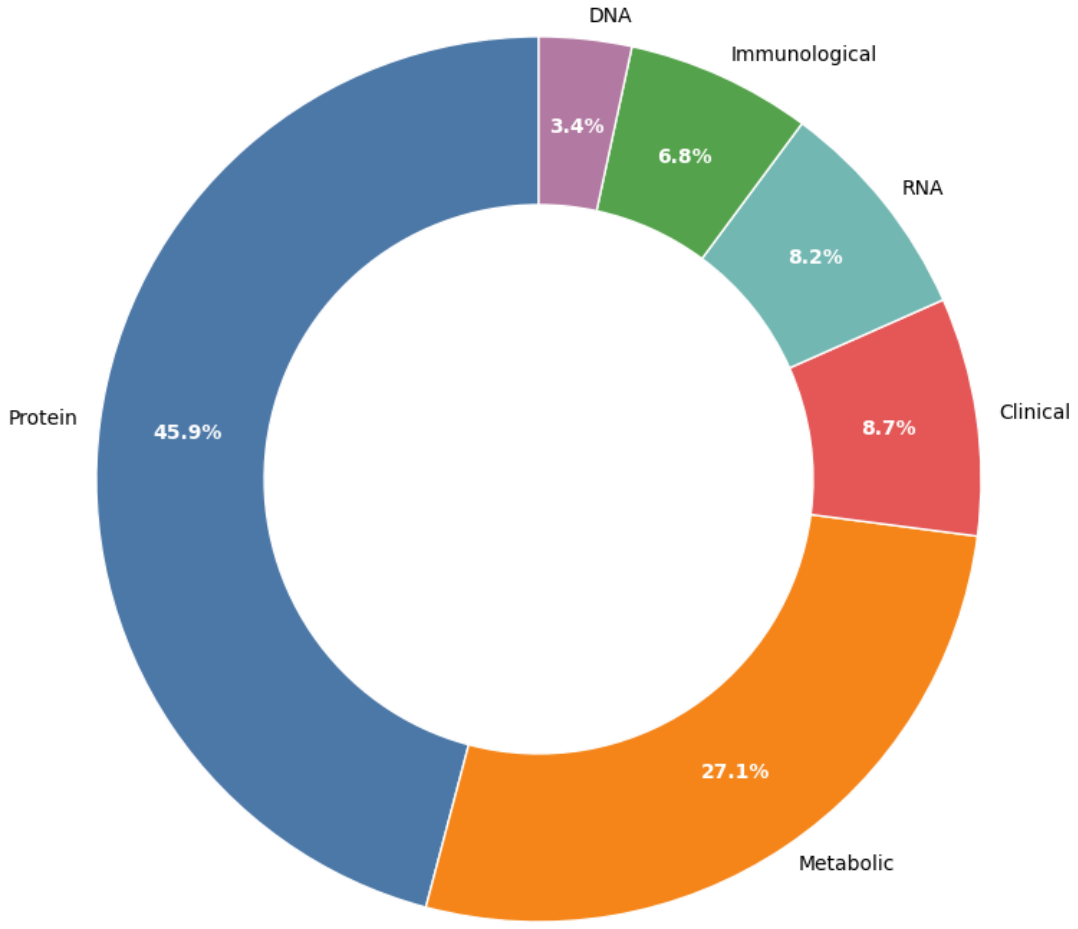


FIGURE 4.2: Showing 207 total biomarkers associated with the metabolic syndrome were curated and later classified into seven major molecular categories.

4.2.2 Disease Categories

All the curated biomarkers were correlated with a wide range of diseases related to metabolic syndrome. The most common disease category was cardiovascular disease (85 biomarker associations) followed by endocrine and metabolic disorders (56). Cancer-related conditions associated with 24 biomarkers, neurological and cognitive disorders associated 10 and liver diseases associated 9.

There were 7 biomarkers associated with kidney diseases and 6 for reproductive disorders and 5 for respiratory disorders (Table 4.2). There were 16 other disease conditions with additional biomarkers. The findings suggest metabolic syndrome markers are linked to various organ systems and disease patterns with cardiovascular disorders being the major clinical phenotype.

TABLE 4.2: Disease categories and respective count of the associated biomarkers

Disease Category	**Associated Biomarkers**
Cardiovascular Diseases	85
Endocrine & Metabolic Disorders	56
Cancer	24
Neurological & Cognitive Disorders	10
Liver Diseases	9
Kidney Diseases	7
Reproductive Disorders	6
Respiratory Disorders	5
Others	16

4.2.3 Sample Sources

MetSMarker was developed with the use of a range of biological samples (Figure 4.3). The most common sample source was serum (103 biomarker records), followed by plasma (67) and blood (33). Other specimen types were urine (13), imaging-based samples (12), saliva (5), platelets (4), cerebrospinal fluid (CSF) (3), leukocytes (3), retina-based samples (3), and others (30).

The abundance of serum- and plasma-derived biomarkers gives rise to the widespread use of these markers in metabolic syndrome research and their value in the identification and validation of biomarkers.

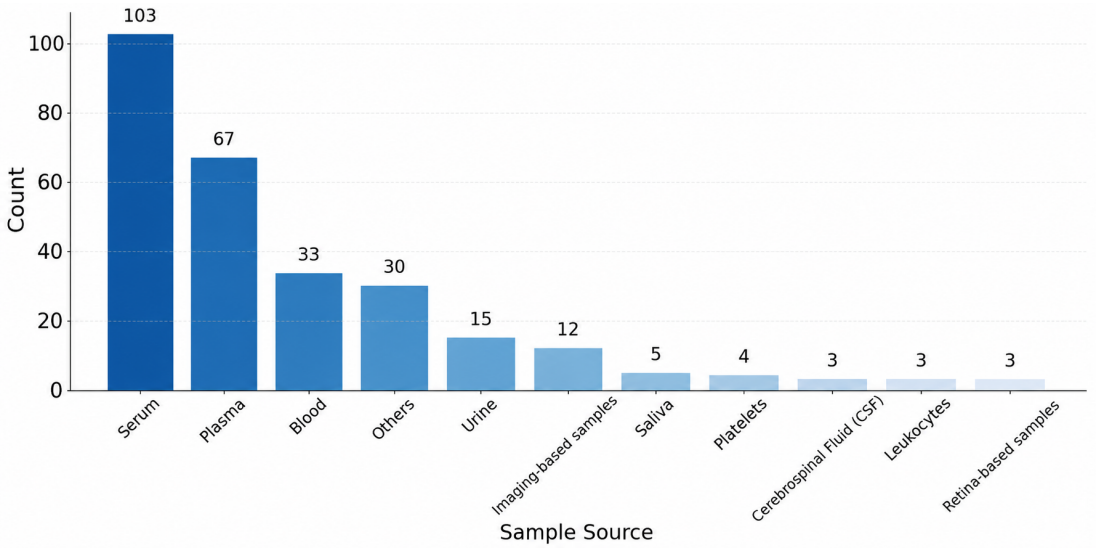


FIGURE 4.3: Showing the biological sample source distribution for the curated biomarkers included in MetSMarker.

4.2.4 Geographic Distribution of Studies

The studies included are representative of countries and regions around the world and demonstrate the global interest in biomarker research. China followed by the United States, Japan, Turkey and the United Kingdom were the countries which had the highest number of publications. Other countries of importance were Italy, Spain, India, Austria, Greece, France, and Taiwan, and several studies were based on multinational and/or international participation, demonstrating a growing trend towards international cooperation in research. The work from developing countries; Bangladesh, Egypt, Indonesia, Romania, Ukraine, Mexico, Brazil, Argentina, South Africa and Tunisia, shows the global impact of biomarker research in different populations and healthcare systems.

Overall, the geographical distribution highlights a high concentration of biomarker research in Asia, Europe and North America, with China being a key player. The relative excess in these regions might be due to increased investment in research, bigger patient numbers or better infrastructure for biomedical research in these areas. Multicenter international studies also indicate growing cooperation among research institutions, resulting in greater external validity and generalizability of biomarker discoveries for other ethnic and clinical populations.

4.2.5 Biomarker Evidence Levels

The curated biomarkers were classified based on the evidence supporting them (Figure 4.4). Tier 3 was the largest category consisting of 172 biomarkers (82.4%). There were 19 biomarkers (9.1%) in Tier 4 and 14 biomarkers (7.2%) in Tier 2.

There was only one biomarker (0.5%) that met the criteria for Tier 1 evidence. The distribution shows that the majority of the biomarkers currently available for metabolic syndrome have moderate evidence levels while comparatively few have been in the highest category of evidence.

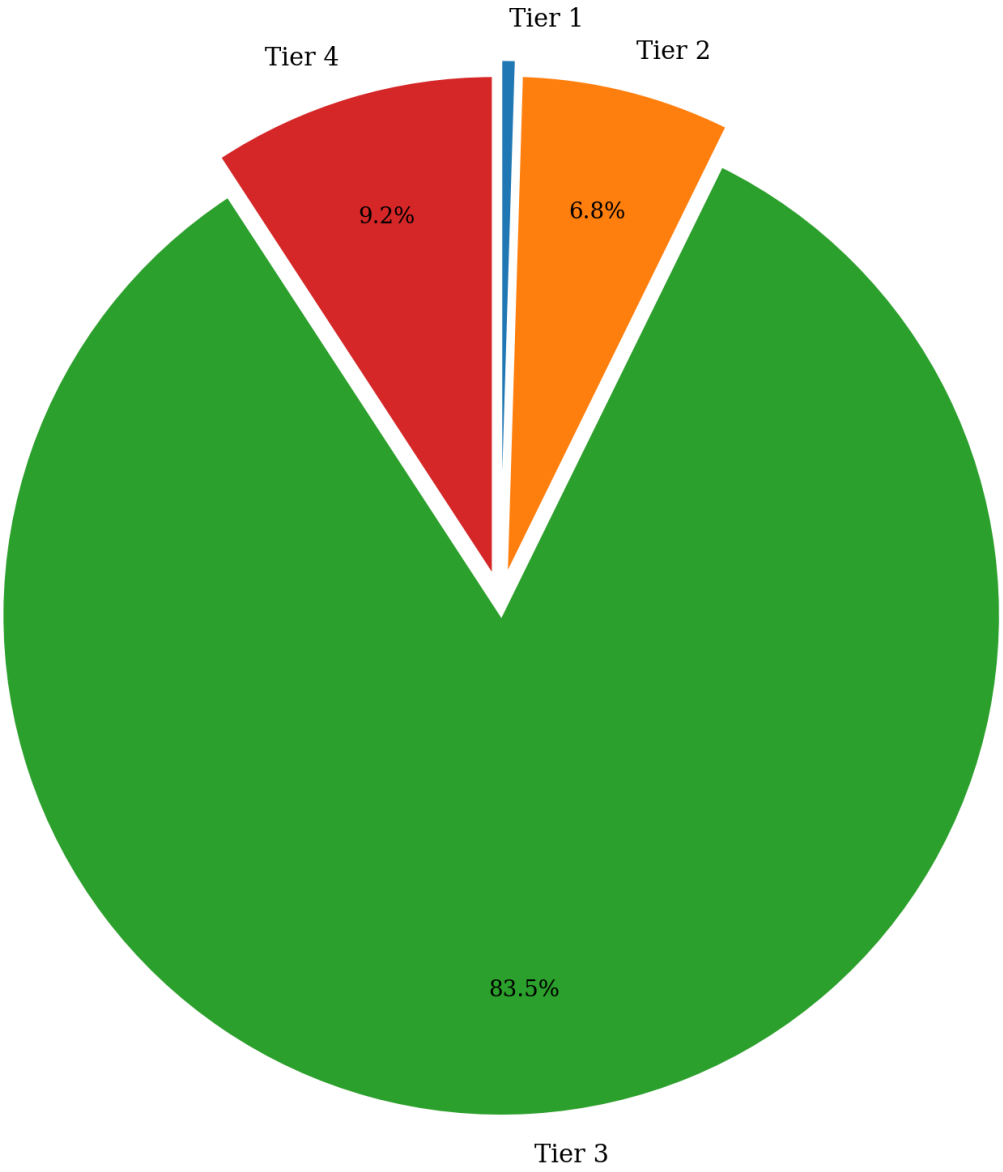


FIGURE 4.4: Showing the level of evidence distribution for the curated biomarkers.

4.2.6 Biomarker Linked Organs

The selected biomarkers correlated with several organs that play a role in the development of metabolic syndrome (Figure 4.5). The most common association of an organ with the 97 biomarkers was the heart, followed by the kidney (55 biomarkers), liver (46 biomarkers), and pancreas (38 biomarkers). Adipose (20), brain (16) and skeletal muscle (14) were other tissues commonly represented. Overall, these data illustrate the multi-system nature of metabolic syndrome and the wide-spread distribution of biomarkers throughout metabolically active organs.

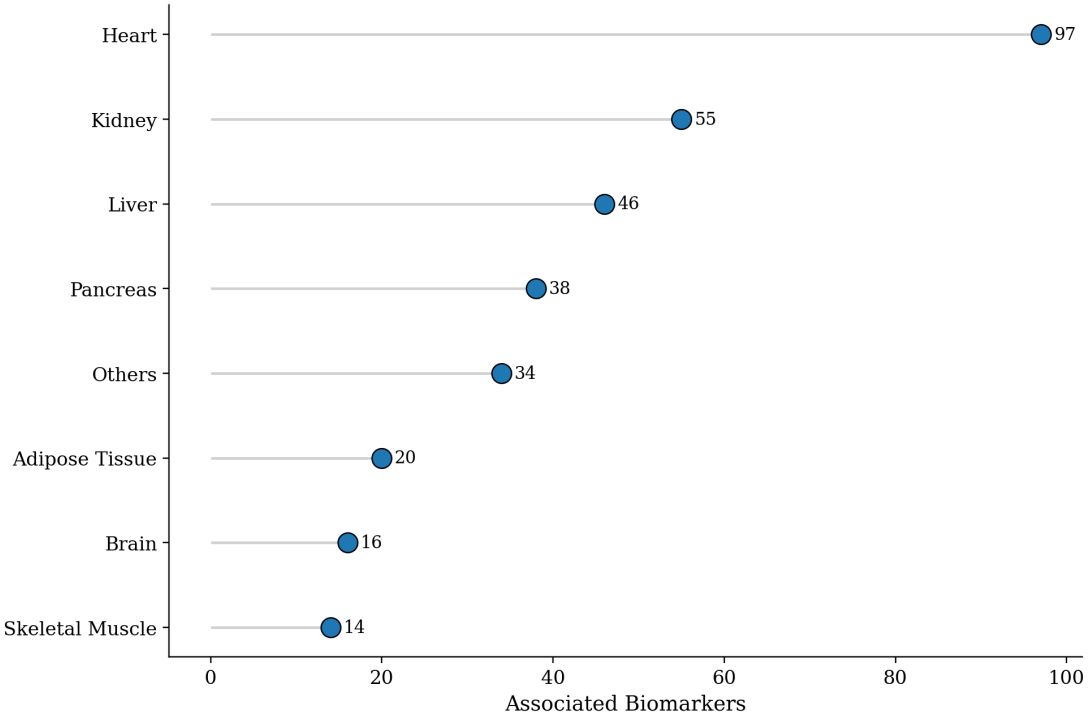


FIGURE 4.5: Showing the association of the curated biomarkers with multiple organs involved in pathophysiology of metabolic syndrome.

4.3 Data Standardization and Annotation

All curated biomarkers have been standardized and annotated thoroughly prior to being incorporated in MetSMarker to make it consistent and interoperable. Harmonized biomarker names were reported in the literature with standardized nomenclature, eliminating inconsistencies due to synonymous names, abbreviations and alternative identifiers.

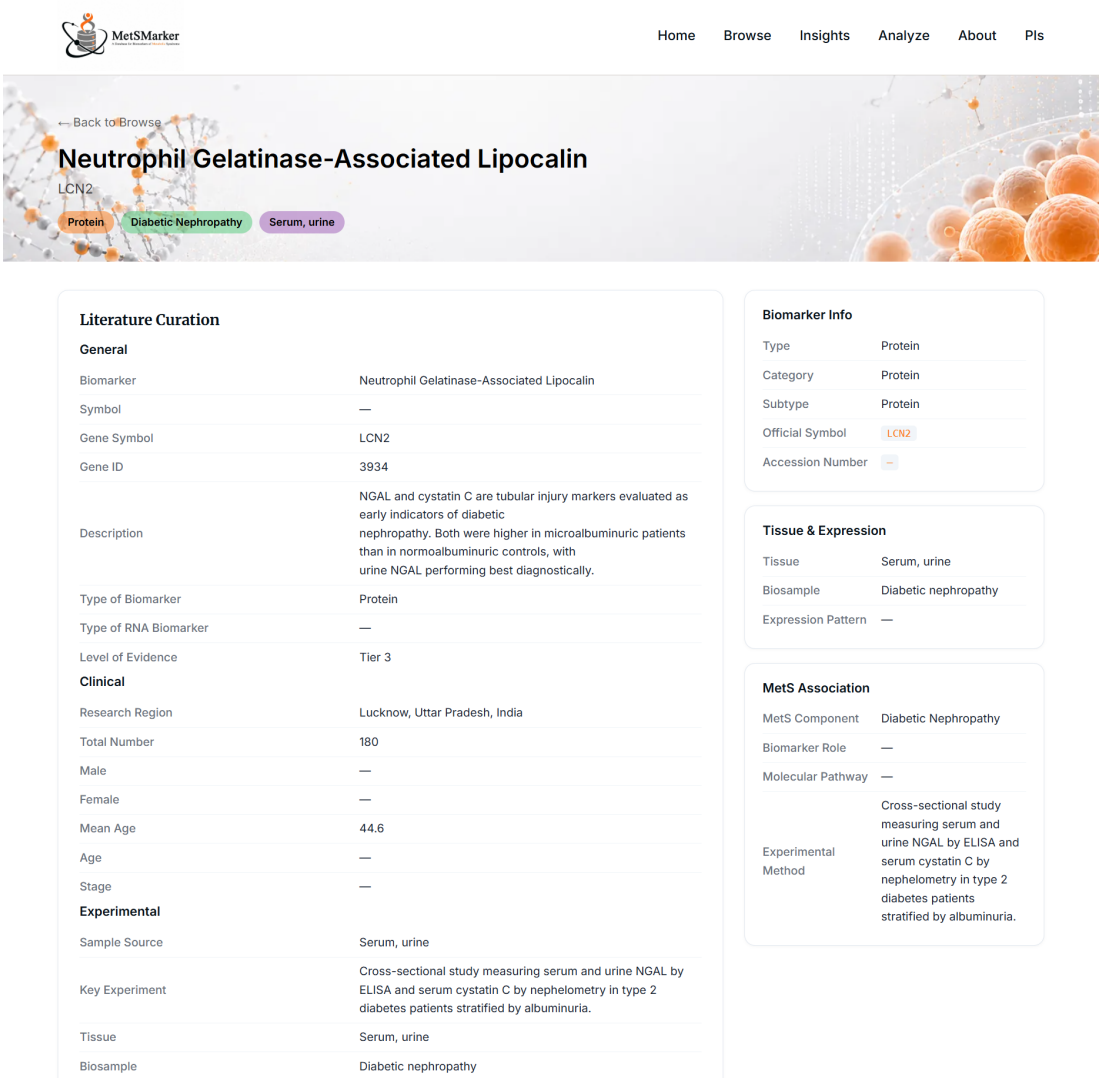


FIGURE 4.6: Biomarkers details showing the comprehensive annotation of the LCN2, Protein, including the classification of biomarker, biological function, molecular pathway, tissue expression, experimental validation, metabolic syndrome association, supporting literature, and external database links.

This standardization provided a standard representation of biomarker records across the platform. Each biomarker record was then annotated with biological and clinical information to enhance data accessibility and to enable the downstream analyses. The annotation framework included biomarker classification, related diseases, source of biological samples, organ association, evidence level and supporting publication information.

Moreover, molecular biomarkers were correlated with external biological repositories such as NCBI Gene ID, which provided standardized information on gene

specific data directly from the platform. The standardized annotation system allows efficient searching, browsing and comparative analysis of articles, and helps minimize the ambiguity of inconsistent terminology for biomarkers used in published articles. Representative annotated records illustrate the ability to combine the use of standardized biomarker nomenclature with curated biological attributes and external resource identifiers, giving a structured and comprehensive view of the metabolic syndrome-associated biomarkers in the platform (Figure 4.6).

4.4 Browse Module of MetSMarker

The Browse module of MetSMarker provides the ability to efficiently browse and retrieve information on biomatics associated with metabolic syndrome available in the database. The module offers several browsers to enable easy access to biomarker data from various biological and clinical viewpoints (Figure 4.7). The database can be searched for terms, diseases, biomarker categories and organs, enabling the researchers to easily locate relevant biomarkers and information.

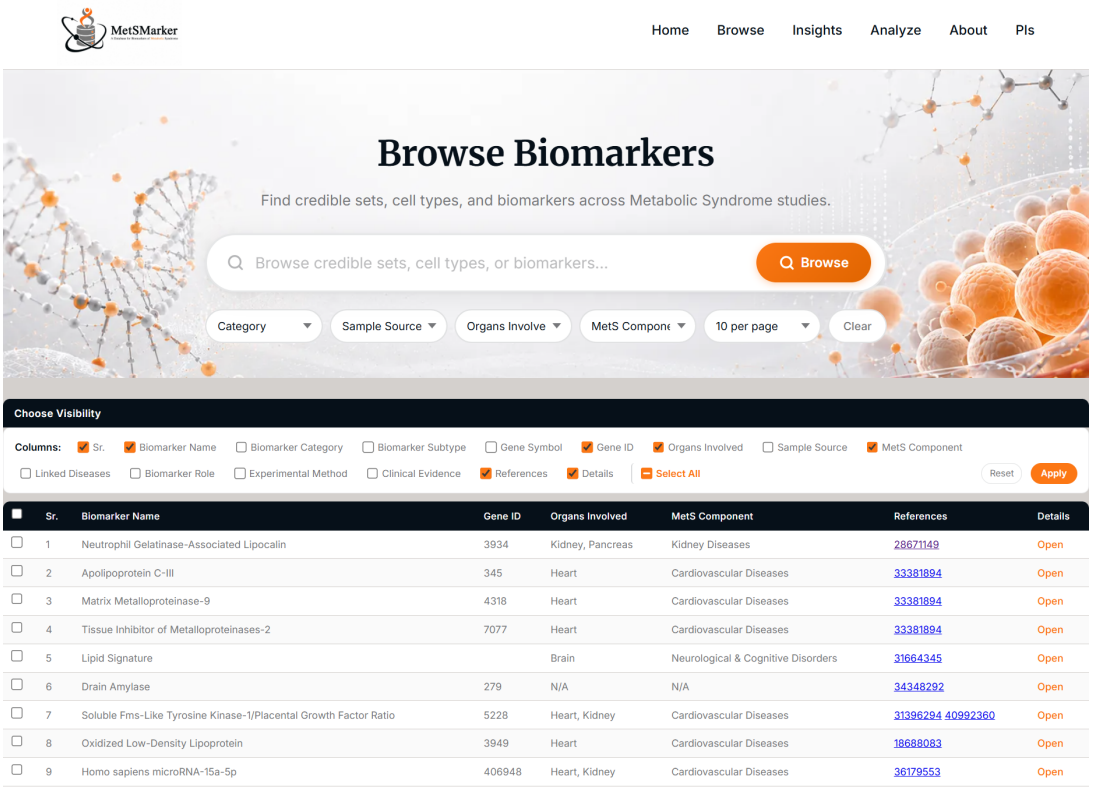


FIGURE 4.7: Interface of Browse module of MetSMarker.

4.4.1 Browse by Term

On this page the user can browse the biomarkers based on the terms selected with which the biomarker data is associated. It works to identify and explore biomarker records associated with the terms in the query and allows users to easily get access to studies and datasets related to specific research topics in the MetSMarker database (Figure 4.8).

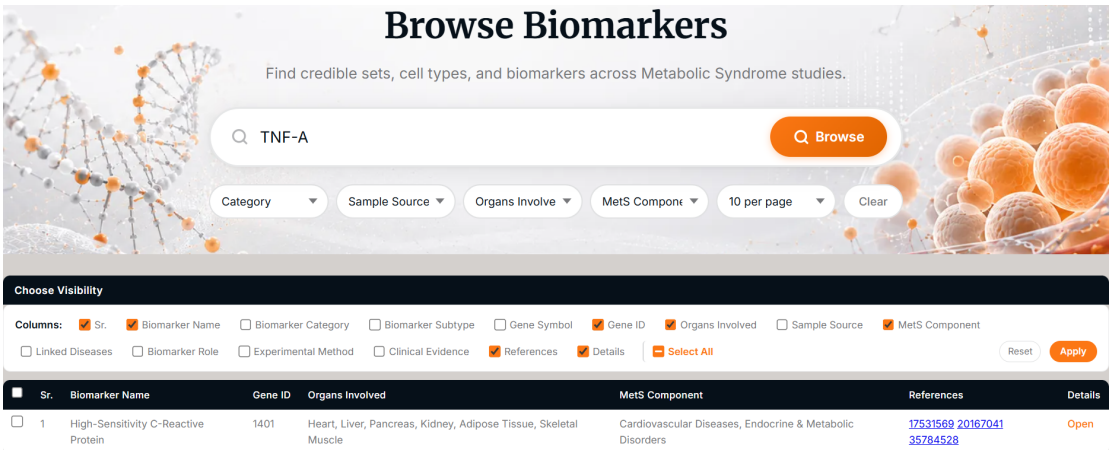


FIGURE 4.8: Browse module of MetSMarker by term e.g. TNF-A.

4.4.2 Browse by Disease

Here you will find an opportunity to navigate through biomarkers by disease associated with metabolic syndrome. Users can discover disease-specific biomarkers and comprehend their role in various pathological situations (Figure 4.9).

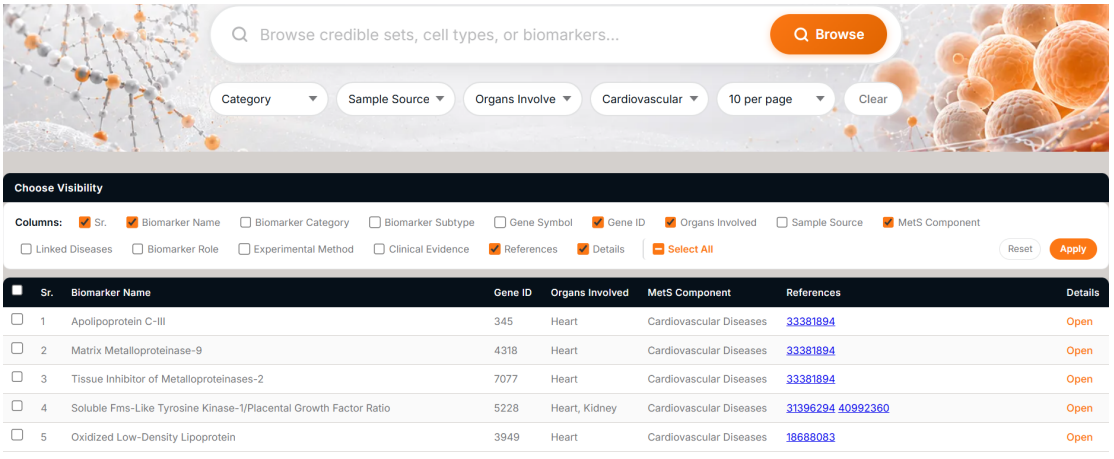


FIGURE 4.9: Browse module of MetSMarker by Disease e.g., CVDs.

4.4.3 Browse by Biomarker Category

The user can browse the database by the different kinds of biomarkers using this page. It aids in the discovery of biomarkers within specific functional or biological classes (Figure 4.10).

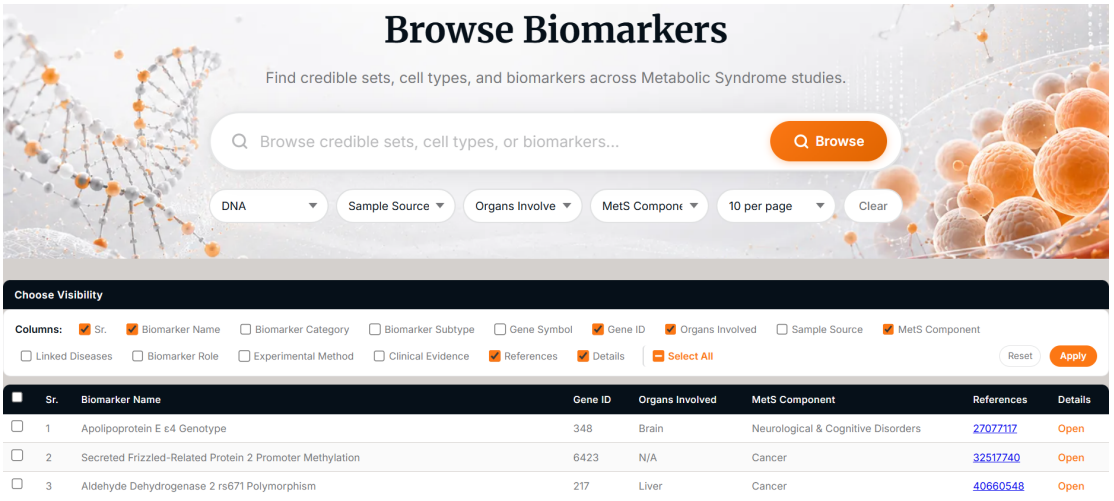


FIGURE 4.10: Browse module of MetSMarker by Category e.g., DNA

4.4.4 Browse by Organ

This page provide with the ability to browse biomarkers by organ that they are expressed or associated with. It provides opportunities to study organ-specific biomarkers of metabolic syndrome and their possible involvement in disease mechanisms (Figure 4.11).

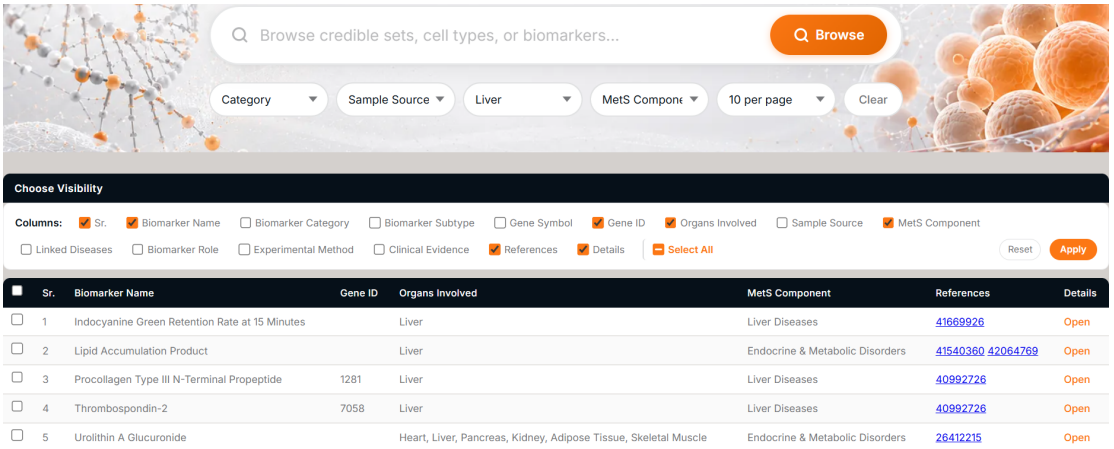


FIGURE 4.11: Interface of Browse module by Organ e.g. Liver

4.5 Insights Module of MetSMarker

MetSMarker's Insights module gives detailed statistical and analytical reports on the contents of the database. This module aims to provide users with insights into biomarker distribution, the selected characteristics of the biomarkers and the prevalence of metabolic syndrome worldwide. This module helps researchers to detect patterns and trends in the database (Figure 4.12).

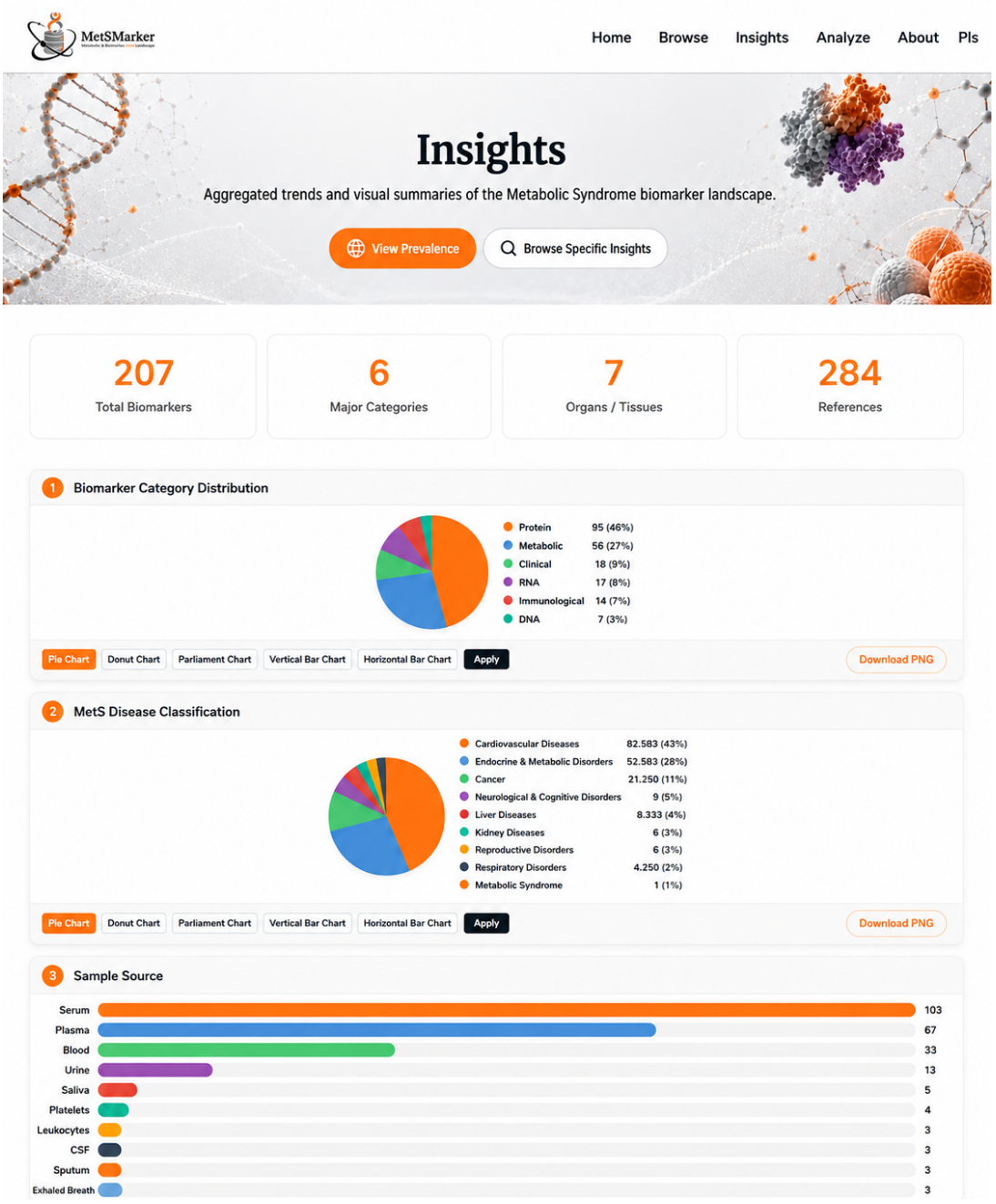


FIGURE 4.12: The interface of insights module of MetSMarker displaying the statistical summaries and visualizations of biomarkers.

4.5.1 Insights of All Database

The overall summary and statistical overview of all biomarkers and datasets in the MetSMarker database is retrieved in this page. It allows users to know the range and contents of the database (Figure 4.13).

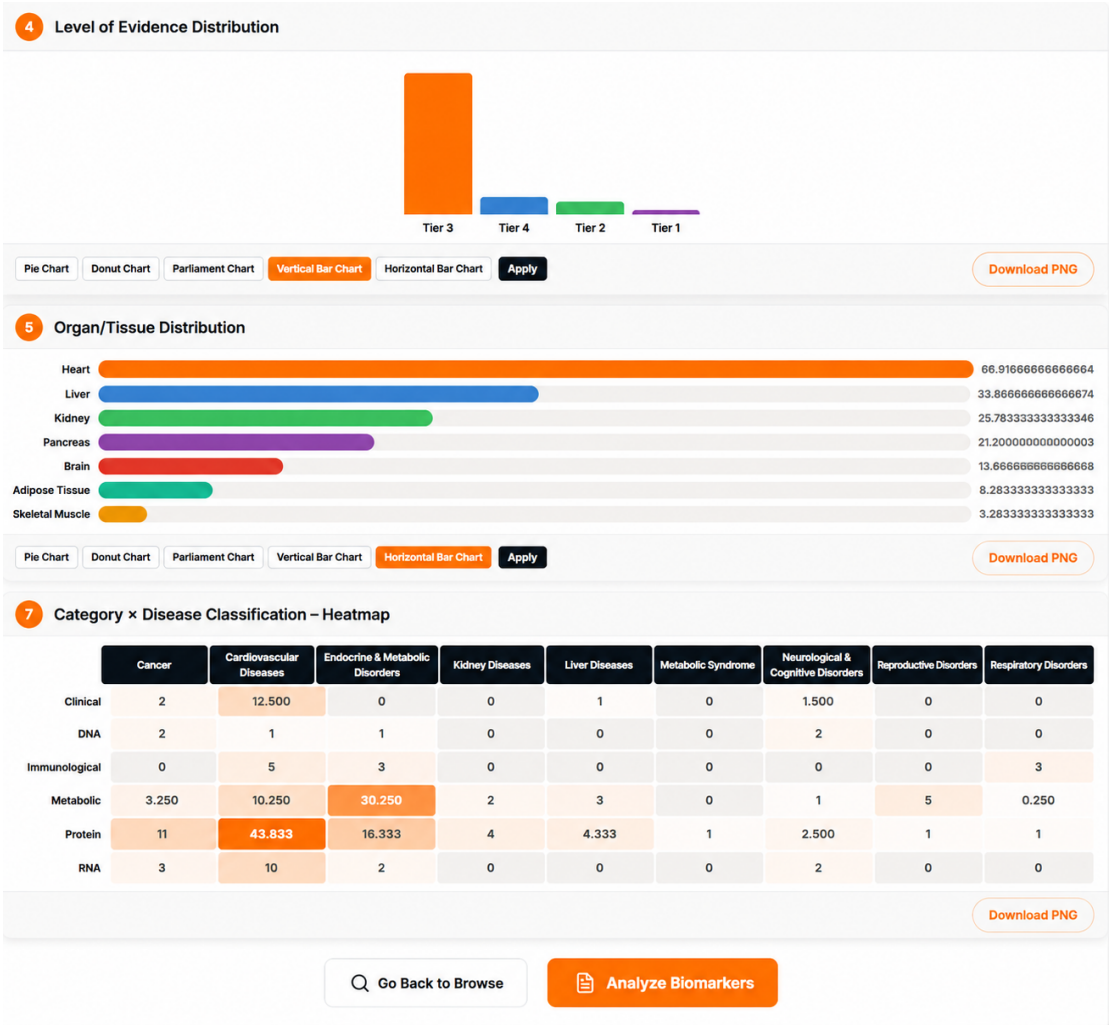


FIGURE 4.13: Insights module of MetSMarker displaying the statistical summaries and visualizations of selected biomarkers, including the category of biomarker, organ/tissue distribution, and the category–MetS component's heatmap.

4.5.2 Insights of Selected Biomarkers

This page contains detailed information and summary statistics of user-selected biomarkers. It assists researchers to examine particular markers and data in more depth (Figure 4.14).

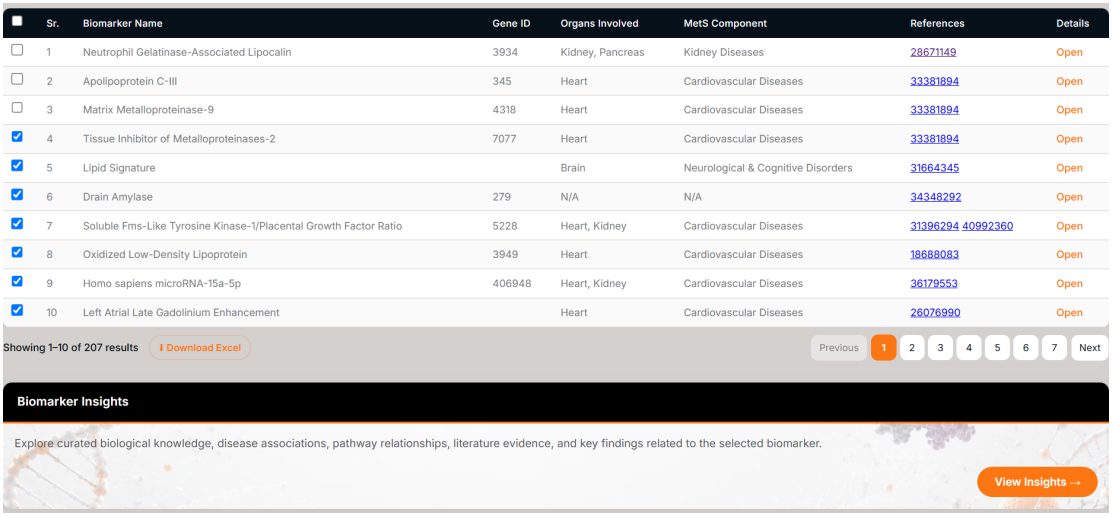


FIGURE 4.14: Insights Module's interface and statistics of selected biomarkers

4.5.3 Prevalence Globally

The following information relates to the global burden of metabolic syndrome. It offers a worldwide picture of the burden of metabolic syndrome and its distribution in populations and region (Figure 4.15).

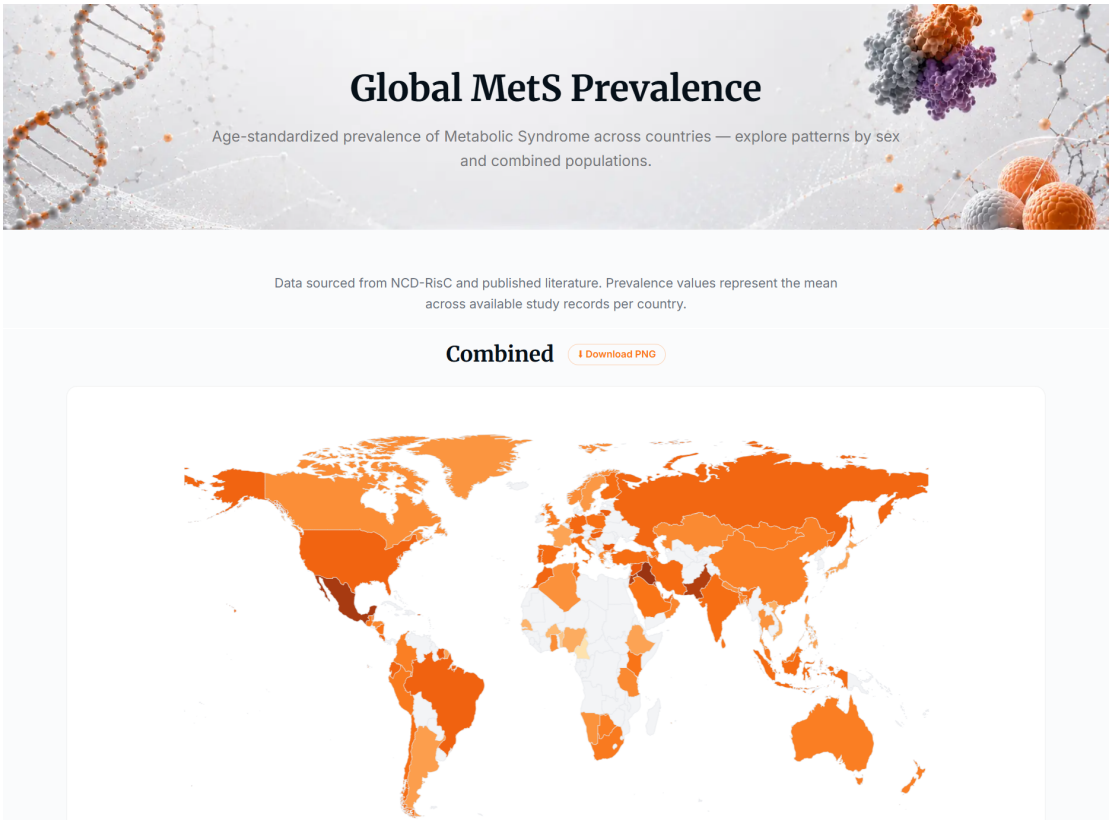


FIGURE 4.15: Global MetS prevalence page represent combined map for male and female

4.6 Analysis Module of MetSMarker

The Analysis module of MetSMarker facilitates comparative and integrative analyses of metabolic syndrome biomarkers. This module supports comparing biomarkers either on one by one basis or in bulk, and cross-omics analyses to uncover interactions and relationships between biomarkers. Results are reported in comparative tables and network diagrams, giving meaningful insights in biomarker associations (Figure 4.16).

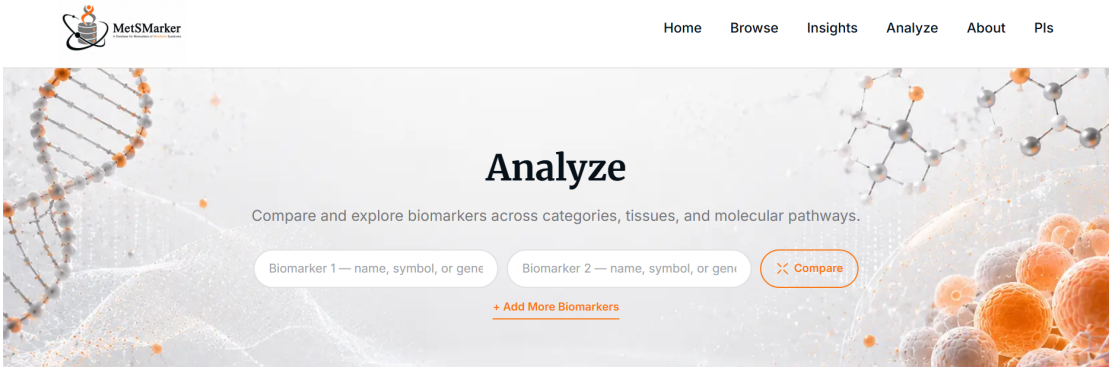


FIGURE 4.16: Interface of Analysis Module of MetSMarker

4.6.1 Comparison of Two Biomarkers

This page enables the user to compare two chosen biomarkers. The results are then presented in a comparative table and user can compare similarities and differences between the selected biomarkers (Figure 4.17).

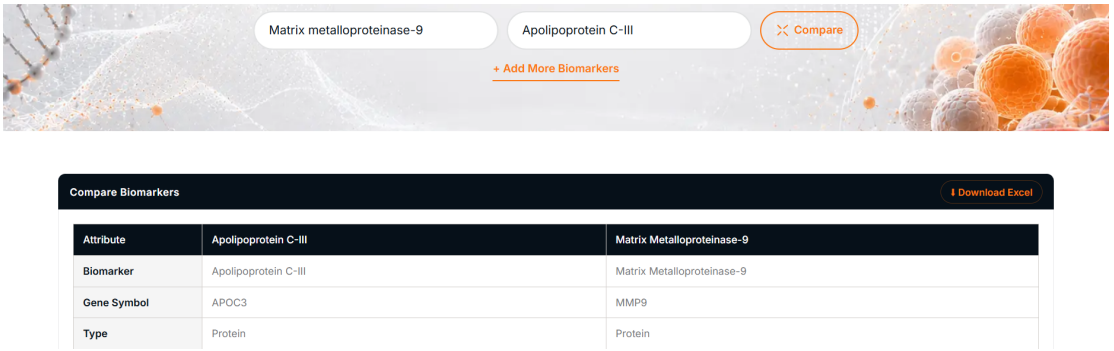
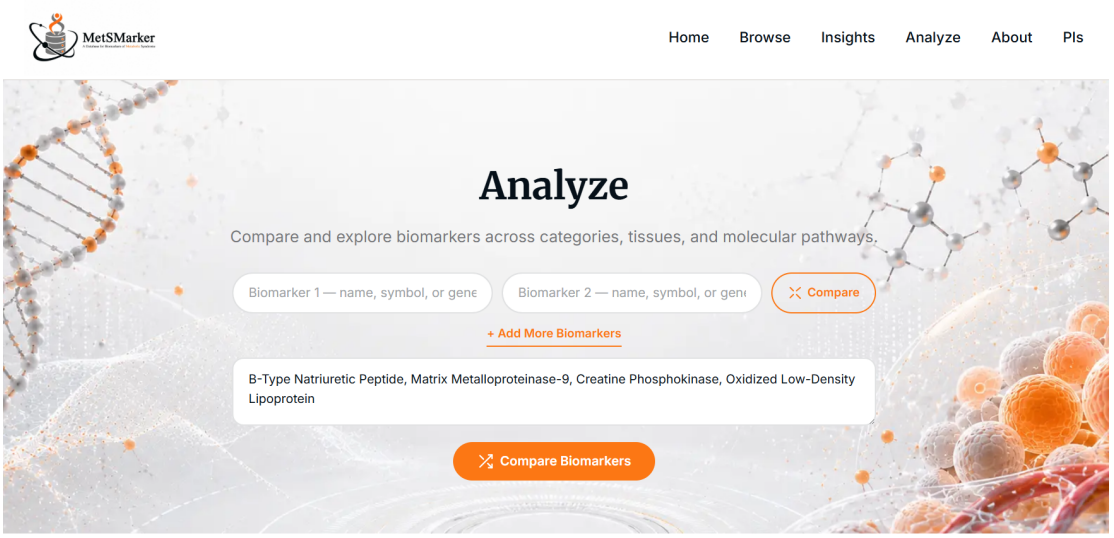


FIGURE 4.17: The interface of Analysis Module of MetSMarker, where users are allowed for comparative analysis of two biomarkers, results are displayed in comparative tables for easy analysis.

4.6.2 Bulk Biomarker Comparison

On this page, user can compare more than one biomarker at a time. The results of analysis are compared and displayed in a comparative table for easy large-scale evaluation and comparison of biomarkers (Figure 4.18).



Compare Biomarkers — Download Excel

Attribute	Matrix Metalloproteinase-9	Oxidized Low-Density Lipoprotein	Soluble Lectin-Like Oxidized Low-Density Lipoprotein Receptor-1	N-Terminal Pro-B-Type Natriure
Biomarker	Matrix Metalloproteinase-9	Oxidized Low-Density Lipoprotein	Soluble Lectin-Like Oxidized Low-Density Lipoprotein Receptor-1	N-Terminal Pro-B-Type Natriure
Gene Symbol	MMP9	LDLR	OLR1	NPPB
Type	Protein	Protein	Protein	Protein
Subtype	Protein	Protein	Protein	Protein
Level of Evidence	Tier 3	—	Tier 3	Tier 3, Tier 2, Tier 2
Research Region	Tunisia	China	Japan	United States (Baltimore, Maryl Netherlands (Leiden), Fortaleza United States, Beijing, China, M International, United Kingdom, International
Source	Plasma	Plasma	Serum	Serum, Plasma, Plasma and ser
Disease Classification	Cardiovascular Diseases	Cardiovascular Diseases	Cardiovascular Diseases	Cardiovascular Diseases, Kidne
Disease Type	Coronary Artery Disease	Coronary Artery Disease	Acute Coronary Syndrome	Pulmonary Hypertension
Sensitivity	—	0.917	—	85, 80.4, 82
Specificity	—	0.925	—	48, 76.9
AUC	0.995	0.975	—	0.75, 0.697
Total Number	757	360	107	98, 80, 153, 6781, 168, 2482, 14
Male	N/A	N/A	71	15, 26, 78
Female	N/A	N/A	36	83, 54, 75
Mean Age	N/A	N/A	64.2	55, 90, 62.5, 87.4
Age	N/A	N/A	Adults	55 Â± 14 years, Nonagenarians, 11.1 years, 45-84 years, >=80 y Adults with chronic dyspnea
Stage	N/A	N/A	N/A	Free of baseline cardiovascular CKD stage 3

FIGURE 4.18: The interface of Analysis Module of MetSMarker, where users are allowed for comparative analysis of bulk biomarkers, results are displayed in comparative tables for large scale analysis and evaluation.

4.6.3 Cross-Omics Analysis

This page enables cross-omics analyses to explore relationships and interactions between biomarkers from different omics layers. Network diagrams are used to visualize the results, giving intuitive understanding of the relationship between biomarkers and the biological association between the biomarkers (Figure 4.19).

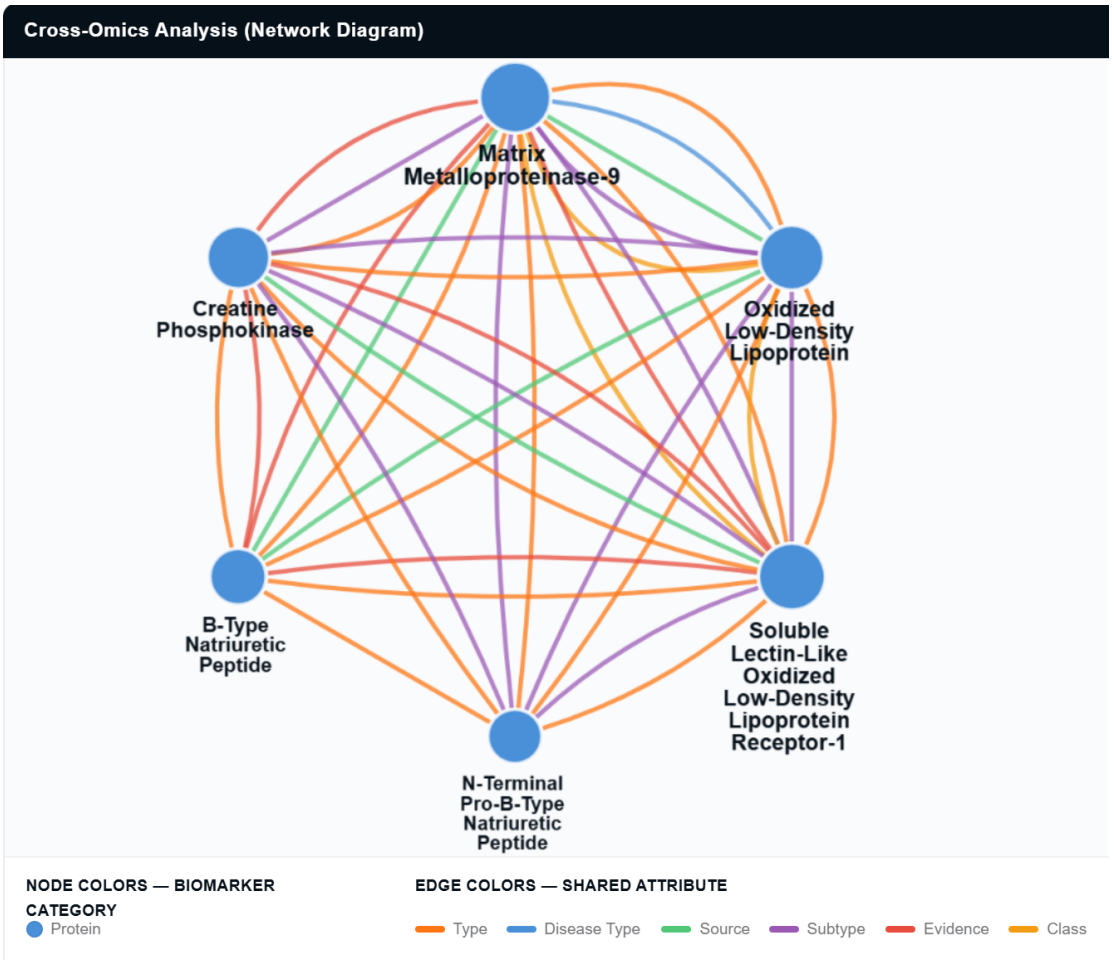


FIGURE 4.19: The Cross-Omics Analysis, results are visualized via network diagrams enabling better understanding of biological associations.

4.7 Database Architecture and Web Interface

4.7.1 Homepage

To make MetSMarker available for efficient exploration and analysis in a user-friendly way, it was implemented in the form of an application in the Web. The

platform features an intuitive interface that allows easy navigation between its key functional components, facilitating users to navigate curated biomarker information in a structured and interactive space (Figure 4.20).

The homepage is the main entry point and links directly to the Browse, Analysis, and Insights modules as well as general information about the platform.

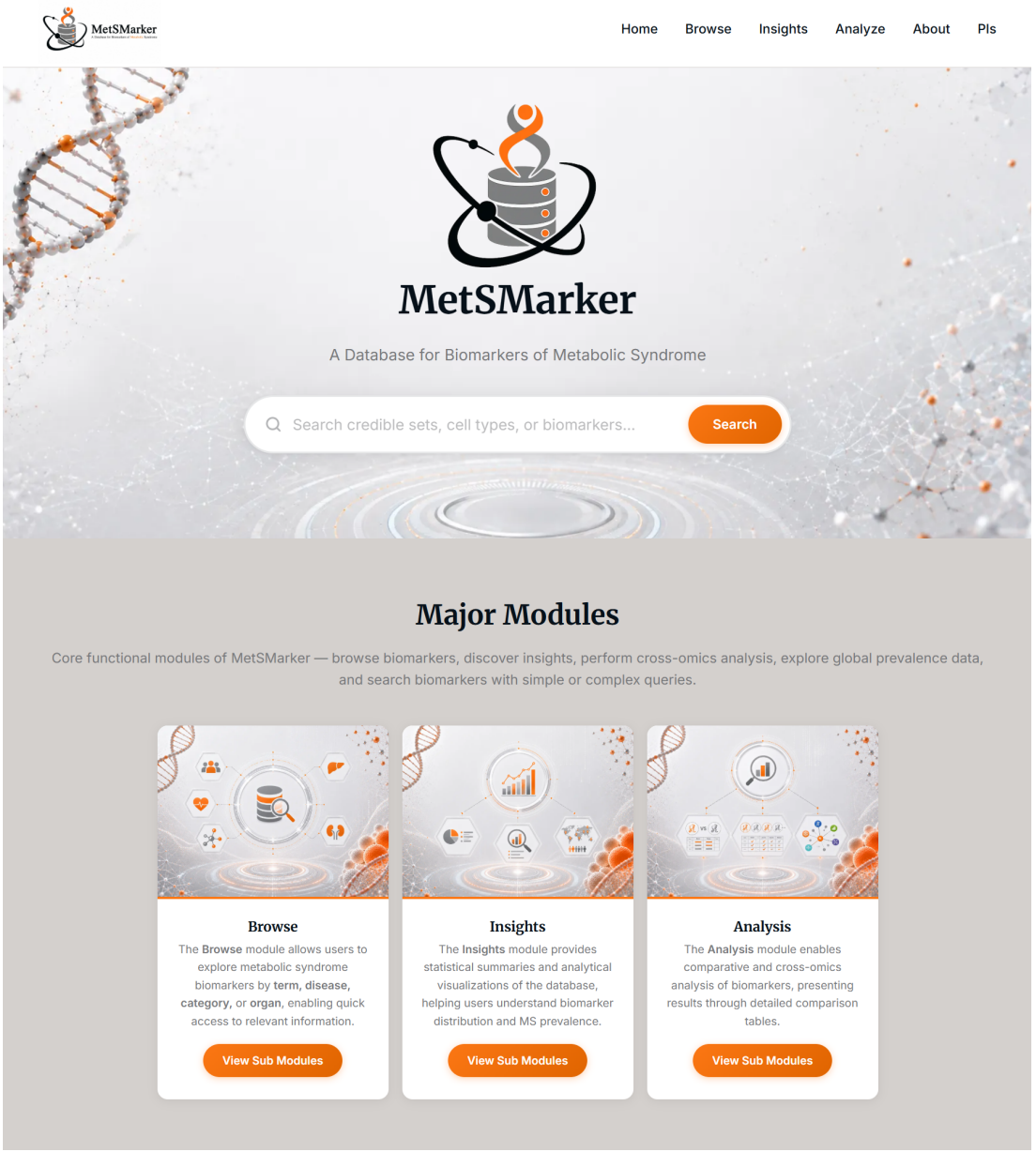


FIGURE 4.20: Home page of MetSMarker, major modules implemented in the MetSMarker: (a) Browse module for the searching as well as exploring metabolic syndrome-associated biomarkers, (b) Insights module specific for visualizing biomarker statistics and their epidemiological information, and (c) Analyze module for the comparative biomarker analysis as well as cross-omics investigation.

4.7.2 Navigation

The Browse module provides a few different ways of accessing biomarker information, such as by browsing by biomarker term, disease, biomarker category, and associated organ. The browsing capabilities enable fast identification of the appropriate biomarker records and exploration from various biological and clinical angles.

The Analysis module offers tools for comparative analysis of biomarkers (Figure 4.20). Users can compare two biomarkers, compare multiple biomarkers in bulk and do cross-omics analyses to gain deeper insight into the relationships between different molecular classes. The created comparative tables and network visualizations enable a detailed examination of the properties of the biomarkers and molecular relations.

The Insights module gives statistical summaries and graphical visualization of the overall composition of the curated collection of biomarkers. This module offers information on the distribution of the biomarkers, selected biomarker characteristics and the worldwide distribution of metabolic syndrome, allowing for a snapshot of the data available and the major trends of the platform.

4.7.3 About Us

In addition to the above functional modules, the platform also includes an About Us page containing the background information on the project, its objectives, as well as the development team. The overall architecture greatly emphasizes simplicity, accessibility, and an efficient navigation, allowing the researchers to explore, compare, and to analyze metabolic syndrome biomarkers through an efficient unified web interface.

Chapter 5

Discussion

5.1 Introduction

Metabolic syndrome (MetS) is a multifactorial disorder that presents with a series of interrelated metabolic abnormalities, such as insulin resistance, central obesity, dyslipidaemia, hypertension and impaired glucose metabolism. Due to the complexity of its molecular mechanisms and clinical symptoms, several biomarkers have been suggested to enhance the diagnosis, prognosis and monitoring of the disease. Despite all the expanding research on biomarkers, however, the evidence is spread out in thousands of scientific publications, making comprehensive interpretation and use in the clinic more challenging. Cross-study comparisons are further complicated and reproducibility is reduced by variations in the names of biomarkers, experimental methodologies, reporting standards, etc. and by molecular classification. Thus, the identification of validated biomarkers, comparison of results between studies and selection of candidate molecules to further investigate are often a major challenge for researchers [43, 44].

The present study aimed to tackle these issues by developing MetSMarker, an integrated, comprehensive web-based knowledge exploration platform that systematically incorporates the experimentally validated molecular biomarkers of metabolic syndrome. The platform was not only meant to be a repository for biomarkers but a tool that would enable biomarker characterization, standardized annotation,

comparative exploration and interpretation of biological evidence. The platform combines DNA, RNA, protein, and epigenetic markers into a single platform, offering researchers a structured environment for research into the molecular complexity of metabolic syndrome, hypothesis generation, and translational research.

The rise of precision medicine has revealed the need to combine diverse biological data and data into a user-friendly computational tool. The key elements in precision medicine are finding consistent molecular indicators that can be used to forecast disease susceptibility, disease progression and therapeutic response. This has made it imperative that curated biomarker resources take on a key and vital role in biomedical research today, by enhancing data access, minimizing data redundancy, and enabling repeatable studies [92]. For metabolic syndrome, a condition in which several interrelated molecular pathways are involved at the same time in causing the disease, an integrated knowledge resource truly adds value—because no one individual biomarker can provide a whole picture of the disease.

The second major contribution of this work is to highlight the manual literature curation process. While automated text-mining techniques have advanced significantly, manually curated information still offers greater reliability by careful examination of the experimental evidence, identification of the biomarkers, molecular classification, and association with disease. The process of standardization on biomarker annotations also contributes to the interoperability of studies and helps reduce inconsistencies that can occur due to different naming conventions or experimental designs. Such harmonization is increasingly known to be a necessary step for meaningful comparisons of biomarkers and subsequent computational analyses [93].

The platform also helps to achieve the platform's overall goal of making molecular discoveries relevant to the clinic. Identification of molecular changes in metabolic syndrome at an early stage could make better risk stratification, individual monitoring and tailoring of therapeutic strategies easier. MetSMarker will create a useful research infrastructure that enables broadly accessible, experimental validation of specific biomarkers; biological annotations; and comparative exploration both for fundamental biomedical research and for future precision medicine efforts.

5.2 Comparison with Existing Databases

Motivation towards developing MetSMarker arose from certain limitations of the existing resources available based on biomarkers. While several web based platforms have been constructed for diseases such as cancer, neurological disorders, cardiovascular diseases, and other chronic diseases, only a limited number of web platforms have been specifically developed to systematically organize experimentally validated molecular biomarkers of diseases associated with metabolic syndrome. Current resources are limited to a single disease, single molecular class, or a single technology (omics) and therefore do not offer a comprehensive picture of the molecular landscape of complex multifactorial diseases. Thus, researchers studying metabolic syndrome are often forced to use multiple independent sources for genetic, transcriptomic, proteomic, and epigenetic data, resulting in increased time spent in data collection, and the increased chance of data that may be interpreted differently by researchers. Because of this, researchers studying metabolic syndrome are often required to use multiple independent sources to get the genetic, transcriptomic, proteomic, and epigenetic data, which increases the time needed to collect such data and the risk of data being interpreted differently by the researchers [43, 44].

One of the most unique and notable features of MetSMarker is the integration of many groups of molecular biomarkers in a structured format. The present platform is not restricted to one biomarker category, but includes all DNA biomarkers, messenger RNAs, microRNAs, long non-coding RNAs, circular RNAs, protein biomarkers and epigenetic regulators, which are all experimentally validated and published in peer-reviewed literature. This multi-layer organization enables to explore molecular interactions at various levels of biological organization, rather than isolated biomolecules. This is especially crucial because metabolic syndrome is a multi-metabolic and multi-inflammatory pathomorphosis that cannot be explained by a single molecular change [6].

One of the other key differences is that the focus is on manual curation and standard annotation. There are many online resources that are based on automated,

literature mining, which is useful for processing large numbers of publications, but may bring in inconsistencies in the identification of the biomarkers, duplicated information, ambiguous gene symbols or missing experimental details. Unlike this, the manual curation strategy, which is used by MetSMarker, involves manually curating each biomarker entry from experimentally validated, peer-reviewed studies and then normalizing each entry against accepted molecular identifiers and biological annotations. This contributes to more accurate data, reduces data duplicateness, and boosts confidence in the retrieved data, making it more appropriate for downstream computational analyses and translational research [93].

The current platform also goes beyond mere information retrieval and also includes knowledge exploration features for comparative analysis of biomarkers. Most resources available today are mostly searchable repositories that allow the user to find a biomarker record, but do not provide many opportunities for comparative investigation or biological interpretation. MetSMarker was developed specifically for interactive analysis and exploration of biomarker characteristics, molecular classifications, disease associations and functional relationships in a single analytical environment. These features enable investigators to explore common biological characteristics, to directly compare multiple biomarkers together and to make new hypotheses about disease processes and future diagnostic signatures.

The other interesting benefit is that of the inclusion of precision medicine ideas in the system structure. In the present era of biomedical research, individualized disease management, that is based on molecular characteristics, is a much increased focus. Therefore, biomarker resources will not only be expected to structure the biological information, but also to aid in personalized diagnosis, risk assessment, and therapeutic decisions [92]. MetSMarker is a systematic organization of experimentally validated biomarkers, along with standardized biological annotations, providing a valuable basis for future precision medicine applications in metabolic syndrome and for biomarker prioritization, patient stratification and translational investigations.

The platform further enhances accessibility and usability by providing a user-friendly web platform specifically designed for efficient biomarker exploration.

With the ever-exponentially increasing scientific publications, efficient retrieval of molecular information is gaining an increasing significance. A well-designed computational platform eliminates the difficulty of literature searches and allows the researcher to find relevant biomarkers in a short amount of time, without having to hand-crawl hundreds of individual papers. This will allow repeatability of research, but also minimise duplication of effort between different laboratories.

However, MetSMarker is not intended to serve as a substitute for the big international molecular repositories that have genome-wide sequencing data, transcriptomic profiles, protein interactions, metabolomic data, or information on pathways. Rather, its main value is the ability to enhance these existing resources with a disease-agnostic, manually curated, experimentally validated knowledge platform that is dedicated to the metabolic syndrome biomarkers. Combining curated literature evidence with a standardized molecular annotation is an essential link between isolated scientific publications and biomedical applications, and is useful for future biomarker discovery, comparative studies and translational research.

5.3 Strengths

MetSMarker's key features are the integration of a multitude of validated molecular markers related to metabolic syndrome within a uniform framework of knowledge, and the structured organization of these markers into a knowledge platform. The information regarding the biomarkers of metabolic syndrome is published in many scientific publications and retrieval and comparative analysis of that information is difficult and time consuming. This makes the platform much more accessible, and less cumbersome to determine which biomarkers may be suitable. This integrated organisation allows molecular evidence to be assessed more efficiently and helps to develop a more comprehensive understanding of the biological complexity of metabolic syndrome [44]. An additional important advantage is that it allows for the simultaneous inclusion of several types of molecules in a single platform.

MetSMarker includes not only DNA-based biomarkers, but also coding and non-coding RNA molecules, protein biomarkers, and epigenetic markers. As metabolic syndrome is the result of a combination of molecular mechanisms, combining various biomarker classes leads to a better representation of disease biology. This multidimensional structure enables the comparison of investigations across molecular levels and makes it possible for the researchers to investigate any relations that might be overlooked if each category of biomarkers is examined independently [6].

Another key benefit of the platform is manual literature curation. Each biomarker featured in MetSMarker was selected from the peer-reviewed, experimental studies, and not solely based on automated literature mining methods. Computational Text-Mining techniques can be used for the rapid extraction of vast amounts of information; however, they can lead to errors in data due to ambiguity in gene names, multiple entries for the same gene, or incomplete experimental description. Manual verification enhances data reliability by ensuring that biomarker identity, molecular classification, disease association, and experimental evidence supporting the data are thoroughly verified before adding the data. This stringent process boosts confidence in the quality of information and its appropriateness for further computational analysis, as well as translation [93].

The platform's usefulness is further bolstered by the adoption of molecular annotations that are standardised. There are often inconsistencies in the nomenclature, identification and reporting of biomarkers, as well as inconsistencies in terms of the molecular classes, between independent studies. MetSMarker enables a harmonisation of these annotations, thanks to the use of internationally recognised biological identifiers and to the use of standardised terminology, and improves the interoperability and the possibility of reproducing these annotations. Standardization also facilitates easier comparative investigations and allows users to interpret information conveyed by biomarkers without being influenced by the variability in reporting between scientific publications [43].

A second advantage is inclusion of knowledge exploration modules which go beyond traditional information retrieval. The platform is not only a repository for biomarker records that can be searched, it is also a tool for comparative analysis

of biomarkers, biological interpretation and interactive exploration of molecular characteristics. The analytical tools help researchers discover common molecular characteristics, determine functional links, and develop new hypotheses about biological functioning. Interactions between multiple biomarkers are especially important in precision medicine research, where the study of multiple molecular markers may be more informative than the study of any single molecular marker.

The platform also offers an intuitive browser interface for easy access to curated biomarker information. A user-friendly organizational structure enables the retrieval of the relevant molecular information by various search and browsing modalities, making the system more accessible to research users of different computational backgrounds. The platform facilitates multidisciplinary research by clinicians, molecular biologists, bioinformaticians and translational researchers by providing easy access to experimentally-validated biomarker data.

Lastly, MetSMarker lays the groundwork for future precision medicine studies of metabolic syndrome. The systematic assembly of validated biomarkers along with the appropriate annotation or standardization is a valuable resource for biomarker prioritisation, diagnostics investigations, computational modelling and hypothesis generation. With the growing trend of molecular medicine to focus on personalized disease management, organised knowledge platforms like MetSMarker can play a substantial role in enhancing biomarker-driven research and enable future translational uses [92].

5.4 Limitations

The results of the present study have to be interpreted with regard to some limitations. The first is that all information on the platform is strictly from published peer-reviewed literature. Therefore, unpublished, invalidated, and unpublished biomarker data was excluded. Such dependence can create publication bias, with biomarkers showing statistically significant results more likely to be reported

than those showing negative or inconclusive results. The next constraint is associated with the manual curation process. While manual extraction can significantly increase the data's accuracy and reliability, it is also very time-consuming and labour-intensive. With the ever increasing amount of biomedical data in the literature, it is essential to update the data on an ongoing basis to ensure that new biomarkers are added in immediately. If the platform is not maintained periodically with newly discovered items, the platform may not offer a complete representation of the discoveries immediately, which may cause a loss of completeness over time, especially with newly discovered items.

The second constraint is associated with the manual curation process. Manual extraction is quite time-consuming and labour-intensive, but remarkably improves the accuracy and reliability of data. As biomedical literature continues to grow, it is necessary to keep up to date with the latest release so that newly discovered biomarkers can be added in as soon as possible. If updates with new finds are not being added constantly to the site, these may not be represented until the next update, and this could diminish the completeness of the website over time.

The present work is mainly concerned with collection of biomarkers, their standardization and knowledge exploration and not experimental validation. The biological relevance of each biomarker relies on the quality of the initial investigations in which they were discovered. Caution should be used when comparing studies because there is variability in study design, sample size, patient characteristics, diagnostic criteria, laboratory methodology or statistical analysis among published studies, which can affect reported biomarker performance.

A drawback is that there is not sufficient full clinical metadata available. There are numerous published studies that yield molecular findings without a description of the severity of the disease, response to treatment, longitudinal follow-up, ethnicity, lifestyle factors or any other medical conditions. The platform currently puts a focus on molecular characterization instead of individual clinical prediction. A richer clinical information would further enhance the translational applicability and enable more robust biomarker stratification.

In addition, while there are standardized annotations that make the platform more interoperable, the platform does not yet include state-of-the-art predictive algorithms that can automatically prioritize biomarkers based on diagnostic accuracy or clinical utility. Likewise, coupling with transcriptomic, proteomic, metabolomic, and clinical cohort data on a large scale is not included in this work. The extra layers of information may further enhance systems-level understanding of metabolic syndrome and prioritize biomarkers.

Lastly, there has not been a thorough external validation process with a number of independent institutions or international user groups. A wider testing of the system by clinicians, researchers and bioinformaticians would be helpful in giving feedback on usability, functionality and real-life applications, and would contribute to further development of the system.

5.5 Future Improvements

There are some opportunities for further development of MetSMarker which could significantly improve its scientific value and translate impact. The development of automatic literature surveillance, based on the use of NLP and text mining with the support of artificial intelligence, is one of the future directions that needs to be explored. This could help identify new published biomarker studies continuously and minimize manual work to keep up with new updates. However, data quality and reliability should be maintained throughout the curation process with expert verification.

Further development of the platform could also include the incorporation of more omics data, such as metabolomics, lipidomics, proteomics, microbiome, and single cell sequencing. The assembly of these complementarily layered molecules would give a more complete systems biology view of metabolic syndrome and allow the study of complex interactions between disparate biological processes.

Another valuable addition would be the addition of longitudinal clinical data such as disease progression, treatment response, therapeutic results and patient

follow-up. The platform would be further enhanced by the inclusion of molecular biomarkers coupled with clinical phenotypes for better risk stratification, prognosis and personalized treatment choices.

The addition of pathway enrichment analysis, protein interaction network visualization, functional annotation and gene ontology information would further enhance the biological interpretation of curated biomarkers. Such analytical capabilities would be useful for exploring molecular mechanisms better and also for identifying potential targets for therapy and pathways that are engaged in the pathogenesis of metabolic syndrome. Biomarkers could also be ranked using machine learning algorithms based on diagnostic accuracy, biological relevance, or predictive performance. Integrated molecular and clinical data can be used to train predictive models that can help researchers select biomarker panels that are more diagnostic than individual biomarkers.

Additional software features that could be developed include application programming interfaces (APIs) for interoperability with other biomedical resources and bioinformatics software. Better data exchange would promote increased use across the research community and enable the incorporation into larger computational workflows. Lastly, an international collaborative curation system with clinicians, molecular biologists and bioinformaticians would facilitate ongoing enhancement and quality control of the platform. Regular updates and external validation, plus community-based expert contributions, would help keep MetSMarker a reliable and up-to-date knowledge resource that can help power biomarker discovery, translational research and future precision medicine efforts in MetS.

Chapter 6

Conclusion

Metabolic syndrome has become one of the most important public health problems worldwide as its prevalence has been rising, and it has been linked to high prevalence of type 2 diabetes mellitus, cardiovascular diseases, chronic kidney disease, metabolic dysfunction associated steatotic liver disease, and other non-communicable diseases. Many molecular biomarkers have been linked to metabolic syndrome, but the results are scattered across the scientific literature, posing significant challenges for systematic retrieval, comparative evaluation and translational application. This fragmentation has hindered the effective identification of validated biomarkers, the ability to compare experimental data and the development of complete molecular signatures that could facilitate precision medicine. The present study aimed at solving these problems by developing a web based knowledge exploration platform called MetSMarker, which systematically integrates the molecular biomarkers for metabolic syndrome, which have been validated experimentally. The platform was developed following a strict workflow that included systematic literature screening, manual curation, standardization of data, biological annotation and web implementation.

MetSMarker was developed to enable knowledge exploration and biological interpretation, rather than the traditional approach of data storage of biomarker information. The new Browse, Analyse, and Insights module allows efficient retrieval of biomarker information, comparative investigations, and study of the total

molecular landscape of metabolic syndrome. These features make the data more accessible as well as facilitate hypothesis generation, prioritization of biomarkers and translational research. A further significant aspect of this work is the focus on manual curation of good quality. All biomarkers included in the platform have been rigorously selected from peer-reviewed and published scientific literature which has been the object of experimental validation. The standardization of the naming of biomarkers, molecular classification, and biological annotations enhance interoperability and reproducibility between studies.

Systematization of experimental validated biomarkers in this study can help to simplify future studies aimed at the development of biomarker panels, molecular risk prediction, patient stratification and the development of personalized treatment for metabolic syndrome. The study has some limitations, however. The site is based on published and peer-reviewed literature only and thus addresses current scientific evidence. Moreover, hand-curated data has the advantage of quality, but is still time-consuming and requires regular updates to add newly published studies.

The present study successfully met the overall objective of systematically curating, standardizing and accessibly making available a knowledge exploration platform for experimentally established molecular biomarkers involved in metabolic syndrome. In the future, with the growth of the genomics, transcriptomics, proteomics, epigenetics, and artificial intelligence in the fields of biomarker research, the resources provided by MetSMarker will become more valuable to support collaborative biomedical studies, to facilitate clinical translation, and to ultimately help achieve better prevention, diagnosis and personalization of metabolic syndrome management.

Bibliography

[1] E. Kassi, P. Pervanidou, G. Kaltsas, and G. Chrousos, "Metabolic syndrome: Definitions and controversies," *BMC Medicine*, vol. 9, 2011.

[2] J. G. Arantes and et al., "Metabolic syndrome: a review and clinical case," in *Development and its applications in scientific knowledge*, 2023.

[3] A. A. Obeidat, M. N. Ahmad, M. A. Ghabashi, A. Y. Alazzeh, S. M. Habib, D. Abu Al-Haijaa, and F. S. Azzeh, "Developmental trends of metabolic syndrome in the past two decades: A narrative review," *Journal of Clinical Medicine*, vol. 14, no. 7, p. 2402, 2025.

[4] Y. Tekalegn and et al., "Metabolic syndrome and its associated factors among type 2 diabetic patients in southwest ethiopia, 2021/2022," 2022.

[5] S. Swarup and et al., "Metabolic syndrome: An updated review on diagnosis and treatment for primary care clinicians," 2024.

[6] G. Fahed, L. Aoun, M. Bou Zerdan, S. Allam, M. Bou Zerdan, Y. Bouferraa, and H. I. Assi, "Metabolic syndrome: Updates on pathophysiology and management in 2021," *International Journal of Molecular Sciences*, vol. 23, no. 2, p. 786, 2022.

[7] I. Lemieux and J.-P. Després, "Metabolic syndrome: Past, present and future," *Nutrients*, vol. 12, no. 11, p. 3501, 2020.

[8] M. Nisar and et al., "Prevalence of undiagnosed metabolic syndrome using three different definitions and identifying associated risk factors among apparently healthy adults in karachi, pakistan," *Archives of Public Health*, 2024.

[9] M. Nwankwo, C. J. Okamkpa, and B. Danborno, "Comparison of diagnostic criteria and prevalence of metabolic syndrome using who, ncep-atp iii, idf and harmonized criteria: A case study from urban southeast nigeria," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 16, p. 102665, 2022.

[10] M. K. Gulov, S. M. Abdullozoda, G. M. Usmanova, and K. K. Kobilov, "Prevalence of metabolic syndrome in tajikistan," *Health care of Tajikistan*, no. 3, 2023.

[11] J. J. Noubiap, J. R. Nansseu, U. F. Nyaga, A. L. Ndoadoumgue, A. T. Ngouo, D. N. Tounouga, J. J. R. Bigna *et al.*, "Worldwide trends in metabolic syndrome from 2000 to 2023: A systematic review and modelling analysis," *Nature Communications*, 2025.

[12] J. P. Sahoo, J. J. Mukherjee, K. O. Lee, and C. M. Khoo, "Epidemiology of metabolic syndrome in south and south-east asia," in *Metabolic Syndrome: From Mechanisms to Interventions*, 2023.

[13] S. O. Adil, M. A. Islam, K. I. Musa, and K. Shafique, "Prevalence of metabolic syndrome among apparently healthy adult population in pakistan: A systematic review and meta-analysis," *Healthcare (Switzerland)*, vol. 11, no. 4, 2023.

[14] Y. Zhang and et al., "Worldwide trends in metabolic syndrome from 2000 to 2023: a systematic review and modelling analysis," *Nature Communications*, 2025.

[15] C. H. Lim and et al., "The prevalence of metabolic syndrome and the associated factors in a multiethnic upper-middle-income country in asia: findings from a nationwide community-based study in 2023," 2025.

[16] N. R. Kandula and et al., "Metabolic syndrome and chronic disease risk in south asian immigrants: A review of prevalence, factors, and interventions," *Healthcare (Basel)*, vol. 11, no. 5, p. 720, 2023.

[17] D. J. Magliano and E. J. Boyko, *IDF Diabetes Atlas, 11th edition.* Brussels: International Diabetes Federation, 2025.

[18] World Obesity Federation, *World Obesity Atlas 2023*. London: World Obesity Federation, 2023.

[19] K. S. Chong, Y. H. Chang, C. T. Yang, C. K. Chou, H. Ou, and S. Kuo, "Longitudinal economic burden of incident complications among metabolic syndrome populations," *Cardiovascular Diabetology*, vol. 23, no. 1, p. 246, 2024.

[20] J. J. Noubiap, J. R. Nansseu, E. Lontchi-Yimagou, and et al., "Geographic distribution of metabolic syndrome and its components in the general adult population: a meta-analysis of global data from 28 million individuals," *Diabetes Research and Clinical Practice*, vol. 188, p. 109924, 2022.

[21] B. M. Hamooya, L. Siame, L. Muchaili, S. K. Masenga, and A. Kirabo, "Metabolic syndrome: Epidemiology, mechanisms, and current therapeutic approaches," *Frontiers in Nutrition*, vol. 12, p. 1661603, 2025.

[22] F. Morgado, A. Valado, J. Metello, and L. Pereira, "Laboratory markers of metabolic syndrome," *Exploration of Cardiology*, vol. 2, no. 3, pp. 114–133, 2024.

[23] S. Ahmad and T. S. Ahluwalia, "The role of genetic and lifestyle factors in metabolic diseases," *Frontiers in Endocrinology*, vol. 10, Jul. 2019.

[24] Y. Wu and et al., "Sedentary time and the risk of metabolic syndrome: a systematic review and dose-response meta-analysis," *Obesity Reviews*, 2022.

[25] M. Godala, M. Krzyżak, D. Maślach, and E. Gaszyńska, "Relationship between dietary behaviors and physical activity and the components of metabolic syndrome: a case-control study," *International Journal of Environmental Research and Public Health*, vol. 19, no. 11, p. 6562, 2022.

[26] H. A. Al Shehri and et al., "Association between preventable risk factors and metabolic syndrome," 2022.

[27] Y. Wang and et al., "Exposure to air pollution and prevalence of metabolic syndrome: a nationwide study in china from 2011 to 2015," 2022.

[28] W. J. Khalil, M. Akeblersane, A. S. Khan, A. S. M. Moin, and A. E. Butler, "Environmental pollution and the risk of developing metabolic disorders: obesity and diabetes," *International Journal of Molecular Sciences*, vol. 24, no. 10, p. 8870, 2023.

[29] P. Rannaud-Bartaire, B. A. Demeneix, and J. B. Fini, "Pressures of the urban environment on the endocrine system: adverse effects and adaptation," 2023.

[30] S. F. V. Wu and et al., "Association between menopause, postmenopausal hormone therapy and metabolic syndrome," *Journal of Clinical Medicine*, vol. 12, no. 13, p. 4435, 2023.

[31] E. Ghadirzadeh, A. Mahmoodi, N. Yousefi, and et al., "Prevalence of metabolic syndrome among iranian postmenopausal females: a systematic review and meta-analysis," *PLoS ONE*, 2025.

[32] U. Zafar, S. Khaliq, H. U. Ahmad, S. Manzoor, and K. P. Lone, "Metabolic syndrome: An update on diagnostic criteria, pathogenesis, and genetic links," *Hormones*, vol. 17, no. 3, pp. 299–313, 2018.

[33] Y. Rochlani, N. V. Pothineni, S. Kovelamudi, and J. L. Mehta, "Metabolic syndrome: Pathophysiology, management, and modulation by natural compounds," *Therapeutic Advances in Cardiovascular Disease*, vol. 11, no. 8, pp. 215–225, 2017.

[34] S. R. Kim and L. O. Lerman, "Diagnostic imaging in the management of patients with metabolic syndrome," *Translational Research*, vol. 194, pp. 1–18, 2018.

[35] S. O'Neill, L. O'Driscoll, and et al., "Systematic review of metabolic syndrome biomarkers: A panel for early detection, management, and risk stratification in the west virginian population," *International Journal of Medical Sciences*, vol. 13, no. 1, pp. 25–38, 2016.

[36] E. Gesteiro, A. Megía, A. Guadalupe-Grau, S. Fernandez-Veledo, J. Vendrell, and M. González-Gross, "Early identification of metabolic syndrome risk: A review of reviews and proposal for defining pre-metabolic syndrome status,"

Nutrition, Metabolism, and Cardiovascular Diseases, vol. 31, no. 9, pp. 2557–2574, 2021.

[37] Y. Cho and S. Lee, "Useful biomarkers of metabolic syndrome," *International Journal of Environmental Research and Public Health*, vol. 19, no. 22, p. 15003, 2022.

[38] P. N. Brandão-Lima, G. B. de Carvalho, T. B. Payolla, F. M. Sarti, and M. M. Rogero, "Circulating microrna related to cardiometabolic risk factors for metabolic syndrome: A systematic review," *Metabolites*, vol. 12, no. 11, p. 1044, 2022.

[39] Y. H. Cho and S. Y. Lee, "Biomarkers in metabolic syndrome," *Advances in Clinical Chemistry*, vol. 111, pp. 101–156, 2022.

[40] F. Åberg, C. D. Byrne, C. J. Pirola, V. Männistö, and S. Sookoian, "Alcohol consumption and metabolic syndrome: Clinical and epidemiological impact on liver disease," *Journal of Hepatology*, vol. 78, no. 1, 2023.

[41] A. A. Hai, S. Iftikhar, S. Latif, F. F. Herekar, S. Javed, and M. J. Patel, "Prevalence of metabolic syndrome in overweight and obese patients and their measurement of neck circumference: A cross-sectional study," *Cureus*, 2019.

[42] B. Mitrovic and et al., "Non-alcoholic fatty liver disease, metabolic syndrome, and type 2 diabetes mellitus: where do we stand today?" *Archives of Medical Science*, vol. 19, no. 4, 2023.

[43] A. Buniello, J. A. L. MacArthur, M. Cerezo, L. W. Harris, J. Hayhurst, C. Malangone, and et al., "The nhgri-ebi gwas catalog of published genome-wide association studies, targeted arrays and summary statistics," *Nucleic Acids Research*, vol. 47, no. D1, pp. D1005–D1012, 2019.

[44] D. S. Wishart, A. C. Guo, E. Oler, and et al., "Hmdb 5.0: The human metabolome database for 2022," *Nucleic Acids Research*, vol. 50, no. D1, pp. D622–D631, 2022.

[45] D. Szklarczyk, R. Kirsch, M. Koutrouli, and et al., "The string database in 2023: Protein–protein association networks and functional enrichment analyses for any sequenced genome of interest," *Nucleic Acids Research*, vol. 51, no. D1, pp. D638–D646, 2023.

[46] M. Kanehisa, M. Furumichi, Y. Sato, M. Kawashima, and M. Ishiguro-Watanabe, "Kegg for taxonomy-based analysis of pathways and genomes," *Nucleic Acids Research*, vol. 51, no. D1, pp. D587–D592, 2023.

[47] United Nations, "Transforming our world: the 2030 agenda for sustainable development," 2015, sDG target 3.4.

[48] C. Y. Ho, J. I. Lee, S. P. Huang, S. C. Chen, and J. H. Geng, "A genome-wide association study of metabolic syndrome in the taiwanese population," *Nutrients*, vol. 16, no. 1, p. 77, 2024.

[49] Y. Wang and et al., "Insulin resistance and metabolic syndrome," 2024.

[50] Y. Chen and et al., "Mechanisms of insulin resistance in metabolic syndrome," 2025.

[51] X. Zhao, X. An, C. Yang, W. Sun, H. Ji, and F. Lian, "The crucial role and mechanism of insulin resistance in metabolic disease," *Frontiers in Endocrinology*, vol. 14, p. 1149239, 2023.

[52] E. Januario, A. Barakat, A. Rajsundar, and et al., "A comprehensive review of pathophysiological link between non-alcoholic fatty liver disease, insulin resistance, and metabolic syndrome," *Cureus*, vol. 16, no. 12, p. e75677, 2024.

[53] Y. Ren, H. Zhao, C. Yin, X. Lan, L. Wu, X. Du, H. R. Griffiths, and D. Gao, "Adipokines, hepatokines and myokines: Focus on their role and molecular mechanisms in adipose tissue inflammation," *Frontiers in Endocrinology*, vol. 13, 2022.

[54] X. Hildebrandt, M. Ibrahim, and N. Peltzer, "Cell death and inflammation during obesity: "know my methods, wat(son)"," *Cell Death & Differentiation*, vol. 30, pp. 279–292, 2023.

[55] M. Del Bas Font and et al., "Adipose tissue inflammation and metabolic syndrome," 2024.

[56] S. K. Masenga, L. S. Kabwe, M. Chakulya, and A. Kirabo, "Mechanisms of oxidative stress in metabolic syndrome," *International Journal of Molecular Sciences*, vol. 24, no. 9, p. 7898, 2023.

[57] H. Lee and P. A. Jose, "Coordinated contribution of nadph oxidase- and mitochondria-derived reactive oxygen species in metabolic syndrome and its implication in renal dysfunction," *Frontiers in Pharmacology*, vol. 12, 2021.

[58] S. Alcover, L. Ramos-Regalado, G. Girón, N. Muñoz-García, and G. Vilahur, "Hdl-cholesterol and triglycerides dynamics: Essential players in metabolic syndrome," *Antioxidants*, vol. 14, no. 4, p. 434, 2025.

[59] F. Hu, Y. Zhang, and Y. Song, "Lipid metabolism, metabolic syndrome, and cancer," in *Lipid Metabolism*. InTech, 2013.

[60] K. K. Farmanfarma and et al., "Is non-high-density lipoprotein associated with metabolic syndrome? a systematic review and meta-analysis," 2022.

[61] A. S. Zahedi, M. Zarkesh, B. Sedaghati-Khayat, M. Hedayati, F. Azizi, and M. S. Daneshpour, "Insulin resistance-related circulating predictive markers in the metabolic syndrome: a systematic review in the iranian population," *Journal of Diabetes and Metabolic Disorders*, vol. 23, no. 1, pp. 199–213, 2023.

[62] D. H. Son, H. S. Ha, H. M. Park, H. Y. Kim, and Y. J. Lee, "New markers in metabolic syndrome," in *Advances in Clinical Chemistry*. Academic Press, 2022, vol. 110, pp. 37–71.

[63] C. E. Kosmas, A. Sourlas, K. Oikonomakis, E. A. Zoumi, A. Papadimitriou, and C. E. Kostara, "Biomarkers of insulin sensitivity/resistance," *Journal of International Medical Research*, 2024.

[64] S. Rana, S. Ali, H. A. Wani, Q. D. Mushtaq, S. Sharma, and M. U. Rehman, "Metabolic syndrome and underlying genetic determinants: A systematic review," *Journal of Diabetes and Metabolic Disorders*, vol. 21, no. 1, pp. 1095–1104, 2022.

[65] Z. Al-Homedi, N. Afify, M. Memon, H. Alsafar, G. Tay, H. F. Jelinek, M. Mousa, N. Abu-Samra, and W. Osman, "Genetic studies of metabolic syndrome in arab populations: A systematic review and meta-analysis," *Frontiers in Genetics*, vol. 12, p. 733746, 2021.

[66] S. Mozafari, M. Ashoori, S. M. Emami Meybodi, R. Solhi, S. R. Mirjalili, A. Dehghani Firoozabadi, and S. Soltani, "Association between apoa5 polymorphisms and susceptibility to metabolic syndrome: a systematic review and meta-analysis," *BMC Genomics*, vol. 25, p. 576, 2024.

[67] E. S. van Walree, I. E. Jansen, N. Y. Bell, J. E. Savage, C. de Leeuw, M. Nieuwdorp, S. van der Sluis, and D. Posthuma, "Disentangling genetic risks for metabolic syndrome," *Diabetes*, vol. 71, no. 11, pp. 2447–2457, 2022.

[68] G. Liu, S. Luo, Y. Lei, J. Wu, Z. Huang, K. Wang, P. Yang, and X. Huang, "A nine-hub-gene signature of metabolic syndrome identified using machine learning algorithms and integrated bioinformatics," *Bioengineered*, vol. 12, no. 1, 2021.

[69] G. Liu, Y. Lei, S. Luo, Z. Huang, C. Chen, K. Wang, P. Yang, and X. Huang, "Microrna expression profile and identification of novel microrna biomarkers for metabolic syndrome," *Bioengineered*, vol. 12, no. 1, 2021.

[70] A. Dandare, M. J. Khan, A. Naeem, and A. Liaquat, "Clinical relevance of circulating non-coding rnas in metabolic diseases: Emphasis on obesity, diabetes, cardiovascular diseases and metabolic syndrome," *Genes & Diseases*, vol. 10, no. 6, pp. 2393–2413, 2023.

[71] M. Rashidmayvan, R. Sahebi, and M. Ghayour-Mobarhan, "Long non-coding rnas: a valuable biomarker for metabolic syndrome," *Molecular Genetics and Genomics*, vol. 297, no. 5, pp. 1169–1183, 2022.

[72] Y. Huang, T. He, J. Zheng *et al.*, "Circular rnas in metabolic health: Bridging the gap between molecular biology and therapy," *Cell Death & Disease*, vol. 17, p. 258, 2026.

[73] K. Mruczyk, A. Cisek-Woźniak, M. Molska, and A. Skoczek-Rubińska, "The role of inflammatory markers in linking metabolic syndrome to cognitive decline in middle-aged women: A focus on tnf-α and il-6," *Metabolites*, vol. 15, no. 3, p. 186, 2025.

[74] F. Osei, A. Block, and P. M. Wippert, "Association of primary allostatic load mediators and metabolic syndrome (mets): A systematic review," *Frontiers in Endocrinology*, vol. 13, p. 946740, 2022.

[75] H. Alwan, V. Aponte Ribero, O. Efthimiou, C. Del Giovane, N. Rodondi, and L. Duntas, "A systematic review and meta-analysis investigating the relationship between metabolic syndrome and the incidence of thyroid diseases," *Journal of Endocrinological Investigation*, 2023.

[76] E. Raya-Cano, R. Molina-Luque, M. Vaquero-Abellán, G. Molina-Recio, R. Jiménez-Mérida, and M. Romero-Saldaña, "Metabolic syndrome and transaminases: systematic review and meta-analysis," *Diabetology & Metabolic Syndrome*, vol. 15, p. 220, 2023.

[77] H. Ahmad, Z. Ahmed, S. Kashif, S. Liaqat, and A. Afreen, "Study of metabolic syndrome indicators in newly diagnosed diabetes mellitus type 2 patients in pakistani population," *Nutrition & Health*, 2022.

[78] M. S. Islam, P. Wei, M. Suzauddula *et al.*, "The interplay of factors in metabolic syndrome: understanding its roots and complexity," *Molecular Medicine*, vol. 30, p. 279, 2024.

[79] J. J. e. a. Heindel, "Metabolism disrupting chemicals and metabolic disorders," *Reproductive Toxicology*, vol. 68, pp. 3–33, 2017.

[80] S. J. Baek, H. J. Ban, S. M. Park, S. Y. Kim, S. Lee, and H. J. Jin, "Genome-wide dna methylation profiling reveals candidate biomarkers and probable molecular mechanism of metabolic syndrome," *Genes & Diseases*, vol. 9, no. 4, pp. 833–836, 2022.

[81] R. Pant, P. Firmal, V. K. Shah, A. Alam, and S. Chattopadhyay, "Epigenetic regulation of adipogenesis in development of metabolic syndrome," *Frontiers in Cell and Developmental Biology*, vol. 8, p. 619888, 2021.

[82] S. Schütte, S. Eberhard, B. Burger, M. Hemmerling, S. Rossol, and J. T. Stahmeyer, "Prevalence of metabolic syndrome: Analysis based on routine statutory health insurance data," 2023.

[83] C. Hu and W. Jia, "Multi-omics profiling: the way toward precision medicine in metabolic diseases," *Journal of Molecular Cell Biology*, vol. 13, no. 8, pp. 576–593, 2021.

[84] J. K. P. Vanamala, V. Sivaramakrishnan, and S. Mummidi, "Editorial: Integrated multi-omic studies of metabolic syndrome, diabetes and insulin-related disorders: mechanisms, biomarkers, and therapeutic targets," *Frontiers in Endocrinology*, vol. 15, p. 1537554, 2025.

[85] E. W. Sayers, E. E. Bolton, J. R. Brister *et al.*, "Database resources of the national center for biotechnology information in 2023," *Nucleic Acids Research*, vol. 51, no. D1, pp. D29–D38, 2023.

[86] F. J. Martin, M. R. Amode, A. Aneja *et al.*, "Ensembl 2023," *Nucleic Acids Research*, vol. 51, no. D1, pp. D933–D941, 2023.

[87] UniProt Consortium, "Uniprot: the universal protein knowledgebase in 2023," *Nucleic Acids Research*, vol. 51, no. D1, pp. D523–D531, 2023.

[88] E. Clough, T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, H. Lee, N. Zhang, N. Serova, L. Wagner, V. Zalunin, A. Kochergin, and A. Soboleva, "Ncbi geo: archive for gene expression and epigenomics data sets: 23-year update," *Nucleic Acids Research*, vol. 52, no. D1, pp. D138–D144, 2024.

[89] J. S. Amberger, C. A. Bocchini, F. Schiettecatte, A. F. Scott, and A. Hamosh, "Omim.org: Online mendelian inheritance in man (omim®), an online catalog of human genes and genetic disorders," *Nucleic Acids Research*, vol. 43, no. D1, pp. D789–D798, 2015.

[90] W. S. Jenkins, C. Richardson, A. Williams, and C. R. Williams-DeVane, "Creating a metabolic syndrome research resource using the national health and nutrition examination survey," *Database (Oxford)*, vol. 2020, p. baaa103, 2020.

[91] J. Piñero, J. M. Ramírez-Anguita, J. Saüch-Pitarch *et al.*, "The disgenet knowledge platform for disease genomics: 2019 update," *Nucleic Acids Research*, vol. 48, no. D1, pp. D845–D855, 2020.

[92] E. A. Ashley, "Towards precision medicine," *Nature Reviews Genetics*, vol. 17, no. 9, pp. 507–522, 2016.

[93] D. J. Rigden and X. M. Fernández, "The 2022 nucleic acids research database issue and the online molecular biology database collection," *Nucleic Acids Research*, vol. 50, no. D1, pp. D1–D10, 2022.