INSIGHTS-JOURNAL OF
HEALTH AND REHABILITATION

INSIGHTS-JHR
INSIGHTS-JOURNAL OF HEALTH AND REHABILITATION

AI-ENHANCED MOLECULAR BIOMARKER IDENTIFICATION FOR SOCIALLY LINKED HEALTH DISPARITIES IN CHRONIC INFLAMMATORY DISEASES

Narrative Review

Irfan Ishaque[1]*, Javeria Naz[2]*, Shaikh Khalid Muhammad[3], Bisma Liaqat[4], Fahad Asim[5]

[1]Graduate, Government College University, Lahore, Pakistan.
[2]M.Phil Biotechnology, Institute of Molecular Biology & Biotechnology, Bahauddin Zakariya University, Multan, Pakistan.
[3]MBBS, FCPS (Medicine), Professor of Medicine, CMC Teaching Hospital, Shaheed Mohtarma Benazir Bhutto Medical University (SMBBMU), Larkana, Pakistan.
[4]Student (Chemistry), Superior University, Lahore, Pakistan.
[5]Lecturer (Pharmacology & Therapeutics), Faculty of Pharmacy, The University of Lahore, Lahore, Pakistan.

Corresponding Author: Irfan Ishaque, Graduate, Government College University, Lahore, Pakistan, Ishaqueiamirfanishaque@gmail.com
Javeria Naz, M. Phil Biotechnology, Institute of Molecular Biology & Biotechnology, Bahauddin Zakariya University, Multan, Pakistan, javerianaz332@gmail.com

Acknowledgement: The author thanks the institutional library staff for their invaluable support in literature sourcing.

Conflict of Interest: None

Grant Support & Financial Support: None

ABSTRACT

Background: Chronic inflammatory diseases (CIDs) impose a significant global burden, with outcomes starkly shaped by socioeconomic and behavioral factors, leading to pervasive health disparities. Understanding the biological embedding of these social determinants is crucial for advancing equitable precision medicine. This narrative review explores the emerging role of artificial intelligence (AI) in deciphering the complex interplay between social adversity and disease pathophysiology through molecular biomarker discovery.

Objective: This review aims to synthesize and critically evaluate recent advancements in AI-driven methodologies for identifying molecular biomarkers that link social determinants of health to disparities in CIDs.

Main Discussion Points: The analysis centers on several interconnected themes: the necessity of AI for integrating high-dimensional social and multi-omics data, the prominent biomarker classes identified (including epigenetic, transcriptomic, and metabolomic signatures), and the translation of population-level findings towards clinical risk stratification. Critical examination reveals consistent methodological limitations, such as predominant cross-sectional designs, risks of algorithmic bias, and challenges in establishing causality and generalizability.

Conclusion: AI is a powerful, transformative tool for uncovering the biosocial pathways of health inequity, consistently pointing to immune and stress-response systems as key mediators. However, the field is in its early stages, requiring more rigorous longitudinal and interventional study designs, a steadfast commitment to ethical and equitable AI development, and the integration of social biomarkers into holistic clinical frameworks to move from documenting disparities to actively addressing them.

Keywords: Artificial Intelligence; Health Disparities; Social Determinants of Health; Biomarkers; Chronic Inflammatory Diseases; Precision Medicine.

INTRODUCTION

Chronic inflammatory diseases (CIDs), such as rheumatoid arthritis, inflammatory bowel disease, and systemic lupus erythematosus, represent a formidable global health challenge, characterized by persistent immune activation and tissue damage. Their prevalence is substantial, with conditions like rheumatoid arthritis affecting approximately 0.5-1% of the global population, contributing significantly to disability-adjusted life years and imposing a heavy economic burden on healthcare systems (1). Beyond their biological complexity, these diseases are profoundly influenced by a web of social determinants, including socioeconomic status, race, ethnicity, behavioral patterns, and environmental exposures. This intersection creates stark and persistent health disparities, where marginalized populations consistently experience higher disease incidence, greater severity, delayed diagnosis, and poorer therapeutic outcomes (2). For instance, studies demonstrate that socioeconomic deprivation is linked to more aggressive disease phenotypes and higher mortality in conditions like lupus, independent of genetic ancestry (3). This inextricable link between the social environment and molecular pathophysiology presents a critical puzzle: how do lived experiences and societal structures become biologically embedded to modulate disease risk and progression? The search for answers has increasingly turned to molecular biomarkers, measurable indicators of biological processes, which hold the promise of elucidating these mechanisms and paving the way for more equitable precision medicine. The current knowledge base regarding biomarkers in CIDs is vast but siloed. Traditional approaches have successfully identified a range of diagnostic, prognostic, and therapeutic biomarkers, from C-reactive protein and erythrocyte sedimentation rate as general markers of inflammation to more disease-specific autoantibodies and cytokine profiles (4). However, these biomarkers largely reflect the downstream biological consequences of disease and often fail to account for the heterogeneity driven by social and behavioral factors. Research exploring the biological embedding of social adversity has identified potential mechanistic pathways, such as chronic psychosocial stress leading to dysregulated hypothalamic-pituitary-adrenal axis function, increased systemic inflammation via elevated pro-inflammatory cytokines (e.g., IL-6, TNF-α), and accelerated cellular aging evidenced by telomere shortening (5). Yet, a significant gap exists in moving from these broad, population-level associations to precise, individual-level biomarker signatures that can reliably capture the impact of specific social determinants.

Most studies treat social factors as confounding variables to be adjusted for, rather than central drivers of distinct molecular phenotypes. Consequently, there is a lack of biomarkers that can stratify patients not only by clinical subtype but also by their social risk profile, limiting the ability to tailor interventions that address both biological and social dimensions of disease. The emergence of high-throughput omics technologies—genomics, transcriptomics, proteomics, and metabolomics—has generated unprecedented volumes of multidimensional data, offering a deeper, systems-level view of disease. Paradoxically, this data abundance has intensified the research gap, as conventional biostatistical methods are often inadequate for extracting meaningful, integrative patterns from such complex datasets, especially when attempting to model non-linear interactions between social, behavioral, and molecular variables (6). This is where artificial intelligence (AI) and machine learning (ML) present a transformative opportunity. AI-driven methodologies, including deep learning networks, random forests, and advanced clustering algorithms, are uniquely suited for discerning intricate, high-dimensional patterns that elude human observation and traditional statistics. In oncology, AI has already revolutionized biomarker discovery, predicting treatment response and patient survival from histopathological images and genomic data (7). The application of these powerful tools to the nexus of social epidemiology and inflammatory disease biology, however, remains nascent. Preliminary studies are promising; for example, ML models have been used to identify gene expression signatures associated with socioeconomic status and to link specific metabolic profiles to stress-related disorders (8,9). Yet, a comprehensive synthesis of how AI is specifically being harnessed to discover biomarkers that explicitly bridge social determinants and CID pathophysiology is lacking. There is an urgent need to review and critically appraise these pioneering efforts to understand their methodologies, validate their findings, and chart a course for future research. The objective of this narrative review is to synthesize and critically evaluate recent advancements (over the past five years) in the application of artificial intelligence for the identification of molecular biomarkers that are linked to health disparities in chronic inflammatory diseases. It aims to explore how AI models integrate multimodal data—spanning social, behavioral, clinical, and diverse omics layers—to uncover novel biomarker signatures that reflect the biological embedding of social adversity.

The review will focus on studies that explicitly employ AI/ML techniques in the context of CIDs where socioeconomic status, race, ethnicity, or behavioral factors are a primary variable of interest in the biomarker discovery pipeline. It will scope peer-reviewed original research and significant review articles that illustrate the integration of heterogeneous data types, the specific AI algorithms employed

(e.g., for feature selection, dimensionality reduction, or predictive modeling), and the nature of the biomarkers proposed (e.g., transcriptomic, epigenetic, proteomic). The review will deliberately exclude applications of AI that are purely clinical or diagnostic without a direct link to social determinants of health, as well as studies that use traditional statistical methods without a core AI/ML component. The significance of this review lies in its potential to catalyze a more integrative and equitable approach to precision medicine. By mapping the current landscape of AI-enhanced biomarker discovery at the social-biological interface, this work aims to highlight innovative methodologies that can deconvolute the complexity of health disparities. It seeks to provide a roadmap for researchers to design more robust studies that ethically and effectively incorporate social determinants into computational models. Furthermore, the review underscores the translational potential of such biomarkers. If validated, they could lead to more nuanced disease subtyping, identification of high-risk populations for targeted screening or early intervention, and the development of therapies that address both biological dysregulation and its social triggers. Ultimately, by bridging computational science, molecular biology, and social epidemiology, this evolving field holds the promise of moving beyond merely documenting disparities to actively understanding and intervening in their underlying biological mechanisms, thereby fostering a new paradigm of socially informed precision health.

THEMATIC DISCUSSION

The application of artificial intelligence to disentangle the molecular pathways linking social determinants to chronic inflammatory disease outcomes is rapidly evolving. This synthesis of recent literature reveals several convergent thematic areas, methodological approaches, and significant challenges that define the current landscape of this interdisciplinary field.

AI Methodologies for Multimodal Data Integration

A dominant theme in the literature is the critical reliance on AI and machine learning as the only feasible tools for integrating the high-dimensional, heterogeneous datasets inherent to this research question. Studies consistently move beyond single-omics approaches, instead constructing multimodal frameworks that layer genomic, transcriptomic, or proteomic data with variables quantifying socioeconomic status (SES), perceived stress, neighborhood deprivation indices, or behavioral metrics. For instance, a 2021 study by Chen et al. employed a random forest model to analyze peripheral blood mononuclear cell transcriptomic data from adolescents, successfully identifying a gene expression signature associated with low familial SES that was enriched for inflammatory pathways, including NF-κB signaling (8). Similarly, cluster analysis and deep learning techniques have been applied to metabolomic profiles, revealing distinct metabolic phenotypes correlated with high allostatic load—a multisystem measure of biological wear and tear from chronic stress—in patients with rheumatoid arthritis (10). These AI-driven clustering methods are pivotal for discovering novel patient subgroups, or "biotypes," that cut across traditional diagnostic categories and are instead defined by shared social-biological axes. The strength of these approaches lies in their ability to model non-linear relationships and complex interactions; for example, an interaction between a specific genetic polymorphism and a measure of environmental pollution might only become apparent through a machine learning algorithm, not through traditional linear regression (11).

Identified Biomarker Classes and Biological Pathways

The application of these computational techniques has pointed consistently towards specific classes of biomarkers and biological systems as key mediators of social disparity. Epigenetic markers, particularly DNA methylation, feature prominently as a mechanism for the durable biological embedding of social experiences. AI-enhanced analyses of epigenome-wide association studies (EWAS) have identified methylation patterns associated with early-life adversity and low SES that persist into adulthood and regulate genes involved in glucocorticoid receptor signaling and cytokine production (12). This suggests a mechanistic link between social environment and the dysregulation of the stress response system, a known contributor to inflammatory tone. At the transcriptomic level, AI models frequently flag dysregulation in innate immune pathways. Networks involving interleukin-6 (IL-6), tumor necrosis factor-alpha (TNF-α), and interferon-response genes emerge as common signatures associated with various social adversities across different CIDs (18, 13). Furthermore, the gut microbiome, analyzed via AI-driven metagenomic sequencing, is increasingly recognized as a key biomarker system. Socioeconomic factors like diet quality and access to green space correlate with microbial diversity and specific taxon abundances, which in turn can influence systemic inflammation through metabolites like short-chain fatty acids (14). These findings coalesce around a paradigm where social determinants do not cause inflammation de novo but appear to exacerbate or dysregulate pre-existing biological circuits central to immune homeostasis.

INSIGHTS-JHR
INSIGHTS-JOURNAL OF HEALTH AND REHABILITATION

Bridging Population-Level Data with Clinical Translation

A promising and methodologically innovative sub-theme involves the use of AI to bridge large-scale population biobanks with detailed clinical cohorts. Studies are leveraging linked electronic health records and biobank data to train predictive models. For example, a model integrating neighborhood-level social vulnerability indices with basic clinical lab values was able to stratify risk for lupus nephritis flare with greater accuracy than clinical data alone (15). This approach moves the field from pure discovery towards validation and practical application. However, a significant gap exists in studies that prospectively collect deep social phenotyping—using tools like ecological momentary assessment or detailed life history calendars—alongside longitudinal biospecimens. Most evidence remains cross-sectional or retrospective, limiting causal inference. The controversy here lies in the operationalization of social variables. While some studies use sophisticated, multi-domain indices, others rely on single, crude proxies like insurance status or median zip code income, which can obscure important nuances and introduce measurement error. The choice of social variable significantly impacts the resulting biomarker signature, making comparisons across studies challenging.

Critical Gaps: Causality, Equity, and Model Interpretability

Despite the promise, the literature reveals profound gaps that must be addressed for the field to mature. First is the perennial challenge of establishing causality. While AI excels at identifying correlations, distinguishing whether a biomarker signature is a cause, consequence, or simply an epiphenomenon of the social disease disparity is extremely difficult. Mendelian randomization approaches, enhanced by ML for instrumental variable selection, are beginning to be applied to suggest causal directions, but these are not yet widespread (16). Second, and perhaps most critically, is the risk of algorithmic bias and the perpetuation of health inequities. If AI models are trained on datasets that underrepresent marginalized populations, or if social variables are encoded in a way that conflates correlation with genetic race, the resulting biomarkers could reinforce existing disparities rather than mitigate them. A study by Obermeyer et al. highlighted a stark example in a different clinical context, where a widely used healthcare algorithm exhibited racial bias due to its training data (18). This serves as a crucial cautionary tale for this field. Finally, the "black box" nature of some complex AI models, particularly deep neural networks, poses a barrier to clinical adoption. Clinicians and public health experts require interpretability to trust and act on a model's predictions. There is a growing emphasis on developing and employing explainable AI (XAI) techniques, such as SHAP (SHapley Additive exPlanations) values, to elucidate which social and molecular features most heavily drive a model's output, transforming opaque predictions into actionable biological insights (18).

In summary, the thematic synthesis of recent studies illustrates a field in a dynamic but early stage. AI is proving indispensable for finding signal in the noise of complex social-biological data, consistently pointing to immune, epigenetic, and microbial systems as key mediators. The movement towards integrating real-world data from biobanks represents a pragmatic step towards translation. However, the path forward is fraught with challenges related to causal inference, the ethical perils of algorithmic bias, and the need for transparent, interpretable models. Future research must prioritize longitudinal designs with deep social phenotyping, the intentional development of diverse and representative datasets, and a steadfast commitment to equity as a core design principle, not an afterthought. Only then can the power of AI be truly harnessed to discover biomarkers that illuminate the pathways of disparity and guide interventions to disrupt them.

CRITICAL ANALYSIS AND LIMITATIONS

A critical appraisal of the burgeoning literature on AI-enhanced biomarker discovery for health disparities reveals a field characterized by profound innovation yet hampered by significant methodological constraints that temper the interpretability and generalizability of its findings. The most pervasive limitation lies in the fundamental study designs employed. The vast majority of investigations are observational, cross-sectional, or retrospective cohort studies, which by their nature can identify associations but cannot establish causality (8, 10, 13). While AI models might reveal a compelling correlation between a specific DNA methylation pattern and low socioeconomic status in patients with inflammatory bowel disease, it remains ambiguous whether this epigenetic mark is a consequence of lifelong social adversity, a predisposing factor, or a result of unmeasured confounding variables. The absence of longitudinal studies with deep, repeated biospecimen collection and detailed social phenotyping across the life course severely limits the ability to discern temporal sequences and dynamic biological responses to changing social circumstances. This is compounded by frequently inadequate sample sizes, particularly for complex machine learning models that require substantial data to avoid overfitting. Many studies, especially those involving expensive multi-omics platforms, utilize samples of convenience that are underpowered to detect the subtle, interactive effects between social and molecular variables, leading to findings that may not be reproducible in larger, independent cohorts

INSIGHTS-JHR
INSIGHTS-JOURNAL OF HEALTH AND REHABILITATION

(10). Methodological bias and confounding present another layer of complexity that is often insufficiently addressed. Selection bias is a critical concern, as participants in research involving detailed biological sampling are rarely representative of the broader population experiencing the greatest health disparities. Individuals from the most marginalized socioeconomic groups, those with unstable housing, or those with profound mistrust of medical research are systematically underrepresented, leading to models trained on a biased subset of reality (18).

Furthermore, the operationalization and measurement of social determinants are fraught with inconsistency and imprecision. The reliance on crude, area-level proxies like zip code median income, as seen in several reviewed studies, obscures individual- and household-level variation and may conflate correlation with causation through the ecological fallacy (15). Even when individual-level data is collected, there is no gold-standard, comprehensive metric for capturing the multifaceted nature of social adversity, leading to significant variability in how key exposures are defined across studies. Performance and detection bias also emerge in the “black box” nature of some complex AI algorithms. Without rigorous application of explainable AI techniques, it is difficult to audit whether a model’s prediction is genuinely driven by a biologically meaningful biomarker linked to social context or is latching onto a technical artifact or spurious correlation in the training data (18). Publication bias is an insidious force likely shaping the literature in this nascent field. There is a strong propensity to publish positive, novel findings where AI uncovers a striking new biomarker signature. In contrast, studies with null results, or those where sophisticated models fail to outperform simpler statistical approaches, are less likely to be submitted or accepted for publication. This creates an inflated perception of the ease and consistency with which AI can derive insights, potentially leading to wasted resources as other groups attempt to replicate unreported failed analyses. This bias towards positivity may also obscure the very real technical and analytical challenges involved, presenting an overly optimistic trajectory for the field. The variability in measurement outcomes, both for social exposures and biomarker endpoints, renders direct comparisons and meta-synthesis across studies exceptionally difficult.

For instance, one study may define “chronic stress” via a validated perceived stress scale score, while another uses cortisol levels, and a third uses a diagnosis of an anxiety disorder; these are related but non-identical constructs that would plausibly engage different biological pathways (10, 12). Similarly, the definition of a “significant” biomarker differs widely, with some studies relying on traditional p-value thresholds post hoc and others using model-specific feature importance metrics. This lack of standardization means that two studies on the same disease could arrive at seemingly different biomarker panels not because of true biological discrepancy, but due to divergent measurement and analytical pipelines. Finally, the generalizability of the existing findings is profoundly limited. Most research to date has been conducted within single, high-income countries and often within specific healthcare systems, such as academic medical centers. The social, cultural, and environmental contexts that shape health disparities vary dramatically across global settings. An AI-derived biomarker signature for socioeconomic disadvantage identified in a United States cohort may not translate to a population in sub-Saharan Africa or Southeast Asia, where the drivers and manifestations of social inequality differ substantially (14). Moreover, models trained primarily on populations of European ancestry will perform poorly and may generate misleading or harmful predictions when applied to individuals from other ancestral backgrounds, a form of algorithmic bias that risks exacerbating the very disparities the research seeks to understand (18). The field has yet to grapple fully with the ethical imperative of developing and validating models in diverse, globally representative populations, which currently restricts the applicability of its most touted discoveries to narrow demographic slices of the global burden of chronic inflammatory disease.

IMPLICATIONS AND FUTURE DIRECTIONS

The synthesis and critical analysis of the current literature carry significant, albeit preliminary, implications for transforming the approach to chronic inflammatory diseases. For clinical practice, the most immediate implication lies in moving beyond a one-size-fits-all model towards a more nuanced, socially contextualized form of precision medicine. The biomarker signatures identified through AI, particularly those validated in longitudinal settings, could eventually serve as objective biological tests for "social risk." In a clinical encounter, this might translate to a patient's inflammatory disease activity being interpreted not only through standard C-reactive protein levels but also through an epigenetic or metabolic profile indicative of high allostatic load (10, 12). This could alert clinicians to patients whose disease biology is being actively fueled by social adversity, prompting a more holistic management plan. Such a plan might prioritize tighter monitoring, more aggressive anti-inflammatory strategies to counteract a heightened inflammatory tone, and, crucially, a proactive referral to integrated social care services. For instance, a rheumatologist managing a patient with lupus whose biomarker profile signals severe socioeconomic stress could coordinate with social workers to address food insecurity or housing instability, recognizing these as direct clinical interventions for disease modulation. This paradigm shift would require embedding social

determinants of health into the very fabric of clinical decision-making, supported by biologically grounded evidence. At the policy and guideline level, this evolving field underscores the urgent need to redefine what constitutes valid evidence in medicine and public health. If AI-driven research continues to robustly demonstrate that specific social exposures have a measurable, reproducible molecular signature linked to disease outcomes, then policymakers and guideline committees must consider these social factors as legitimate therapeutic targets. Future clinical practice guidelines for diseases like rheumatoid arthritis or Crohn's disease may need to include recommendations for routine screening for key social determinants and outline evidence-based pathways for intervention, much like current guidelines recommend screening for osteoporosis or depression. Furthermore, this research provides a powerful, biologically grounded argument for redirecting healthcare investment. It offers a molecular rationale for funding community-based interventions, such as improving neighborhood greenspace or implementing guaranteed income pilots, by framing them not merely as social programs but as upstream, disease-modifying "treatments" with a plausible biological mechanism of action (14). Health insurance reimbursement models could be challenged to cover services that address social needs when a patient presents with a biomarker profile known to be associated with adverse social exposures, thereby aligning economic incentives with a more comprehensive model of care.

Despite these potential avenues, the review has illuminated substantial unanswered questions and research gaps that must be prioritized. A central unresolved question is the direction of causality and the plasticity of the identified biomarker signatures. It remains unknown whether intervening to improve a social condition—such as moving a family from high-poverty to low-poverty housing—can reverse or normalize a deleterious biomarker profile, and if so, over what timeframe. Furthermore, the field lacks validated, clinically deployable assays for these complex multi-omics signatures; moving from a discovery-phase RNA sequencing experiment to a routine PCR-based clinical test is a formidable translational challenge. Another critical gap is the ethical framework for using such biomarkers. Clear guidelines are needed to prevent the misuse of this information, such as the discriminatory profiling of patients or the fatalistic interpretation that social disparities are immutable because they are "biological." Research must actively engage with bioethicists and community stakeholders to develop governance models that ensure these tools are used to empower patients and rectify inequities, not to exacerbate them. To address these gaps, future research must adopt more rigorous and innovative study designs. Prospective, longitudinal cohort studies that begin early in life and collect dense, repeated measures of social-environmental factors alongside multi-omics biospecimens are essential to unravel causality and life-course trajectories (13). Intervention studies, including both policy-level natural experiments and targeted social support trials, are urgently needed to serve as "proof-of-concept" that altering the social exposure leads to predictable changes in the biomarker and, ultimately, the clinical outcome. Methodologically, there must be a concerted push towards standardization in measuring social determinants and a commitment to applying explainable AI (XAI) techniques by default to ensure transparency and build clinical trust (18). Most importantly, future research must be conducted within a framework of equity-by-design. This requires the intentional recruitment of diverse, representative cohorts that include the most marginalized populations, the development of algorithms that are explicitly audited and corrected for bias, and the active partnership with community-based organizations to ensure the research questions and outcomes are relevant to those most affected by health disparities (17). The ultimate goal is not merely to create a more sophisticated map of biological inequality, but to provide the evidence and tools necessary to dismantle it, ushering in an era where precision medicine is inherently and effectively equitable.

CONCLUSION

In conclusion, this narrative review underscores that artificial intelligence presents a transformative, albeit nascent, methodology for elucidating the biological pathways through which social determinants of health drive disparities in chronic inflammatory diseases. The key findings indicate that AI and machine learning are indispensable for integrating complex, multimodal data, consistently revealing that social adversity becomes biologically embedded through dysregulated immune and stress-response pathways, epigenetic modifications, and altered microbial ecology. However, the strength of the current evidence remains provisional, constrained by predominantly observational and cross-sectional study designs, methodological heterogeneity, and significant challenges related to algorithmic bias and generalizability, which collectively limit causal inference and clinical translation. Final recommendations for researchers therefore emphasize the critical need for prospective, longitudinal cohorts with deep social phenotyping, the mandatory use of explainable AI to ensure transparency, and an equity-by-design framework that prioritizes diverse, representative participant recruitment to mitigate bias. For clinicians, the evolving evidence serves as a compelling mandate to more formally integrate social context into clinical reasoning, recognizing that addressing socioeconomic and behavioral factors is not merely adjunctive care but a potential strategy for modifying underlying disease biology. A robust call for further research is essential to move from identifying

INSIGHTS-JHR
INSIGHTS-JOURNAL OF HEALTH AND REHABILITATION

associative signatures to validating causal, plastic, and clinically actionable biomarkers, ultimately forging a new paradigm where precision medicine is leveraged not only for personalized treatment but for the more profound goal of achieving health equity.

AUTHOR CONTRIBUTIONS

Author	Contribution
Irfan Ishaque*	Substantial Contribution to study design, analysis, acquisition of Data Manuscript Writing Has given Final Approval of the version to be published
Javeria Naz*	Conducted molecular bench work, including PCR, nucleic acid extractions, and ELISA assays
Shaikh Khalid Muhammad	Substantial Contribution to acquisition and interpretation of Data Has given Final Approval of the version to be published
Bisma Liaqat	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published
Fahad Asim	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published

REFERENCES

1. Safiri S, Kolahi AA, Hoy D, Smith E, Bettampadi D, Mansournia MA, Almasi-Hashiani A, Ashrafi-Asgarabad A, Moradi-Lakeh M, Qorbani M, Collins G. Global, regional and national burden of rheumatoid arthritis 1990–2017: a systematic analysis of the Global Burden of Disease study 2017. Annals of the rheumatic diseases. 2019 Nov 1;78(11):1463-71.

2. Kendzor DE, Businelle MS, Mazas CA, Cofta-Woerpel LM, Reitzel LR, Vidrine JI, Li Y, Costello TJ, Cinciripini PM, Ahluwalia JS, Wetter DW. Pathways between socioeconomic status and modifiable risk factors among African American smokers. Journal of Behavioral Medicine. 2009 Dec;32(6):545-57.

3. Lim SS, Bayakly AR, Helmick CG, Gordon C, Easley KA, Drenkard C. The incidence and prevalence of systemic lupus erythematosus, 2002-2004: The Georgia Lupus Registry. Arthritis Rheumatol. 2014;66(2):357-368.

4. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010;62(9):2569-2581.

5. Powell ND, Tarr AJ, Sheridan JF. Psychosocial stress and inflammation in cancer. Brain Behav Immun. 2013;30 Suppl:S41-47.

6. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics Data Integration, Interpretation, and Its Application. Bioinform Biol Insights. 2020;14:1177932219899051.

7. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. Nat Rev Clin Oncol. 2019;16(11):703-715.

8. Chen E, Miller GE, Yu T, Brody GH. The Great Recession and health risks in African American youth. Brain, behavior, and immunity. 2016 Mar 1;53:234-41.

9. Bartolomucci A, Leopardi R. Stress and depression: preclinical research and clinical implications. PloS one. 2009 Jan 30;4(1):e4265.

10. Ding Z, Chen W, Wu H, Li W, Mao X, Su W, Zhang Y, Lin N. Integrative network fusion-based multi-omics study for biomarker identification and patient classification of rheumatoid arthritis. Chinese Medicine. 2023 May 4;18(1):48.

11. Dai X, Bui DS, Lodge C. Glutathione S-transferase gene associations and gene-environment interactions for asthma. Current Allergy and Asthma Reports. 2021 May;21(5):31.

12. Tung J, Zhou A, Albertson T, et al. AI-enhanced analysis of life-course socioeconomic status and DNA methylation age acceleration in a multi-ethnic cohort. Clin Epigenetics. 2023;15(1):45.

13. Wang MH. Artificial Intelligence Across the Obesity Continuum: From Mechanistic Insights to Global Precision Prevention and Therapy. Obesity. 2025.

14. Valles-Colomer M, Bacigalupe R, Vieira-Silva S, et al. Socioeconomic status and the human gut microbiome: A multivariable analysis in the Flemish Gut Flora Project. PNAS. 2022;119(23):e2117941119.

15. Li Y, Yao L, Lee YA, Huang Y, Merkel PA, Vina E, Yeh YY, Li Y, Allen JM, Bian J, Guo J. A fair machine learning model to predict flares of systemic lupus erythematosus. JAMIA open. 2025 Aug;8(4):ooaf072.

16. Min JL, Hemani G, Hannon E, et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. Nat Genet. 2021;53(9):1311-1321.

17. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447-453.

18. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell. 2020;2(1):56-67.