The Exposome Interpreter: A Multi-Modal Framework for Personalized Autoimmune Care

Kritika Chugh
Valence Wellbeing
Sunnyvale, CA 94086
kritikachugh2@gmail.com

Abstract

The exposome, the totality of an individual's environmental exposures throughout their lifetime is estimated to account for up to 70% of autoimmune disease risk. Despite this significant contribution, the systematic identification of patient-specific environmental triggers remains an intractable challenge in clinical practice. This translational gap arises from the difficulty of synthesizing vast, heterogeneous data sources: semi-structured clinical lab reports, patient product usage history, and the exponentially growing corpus of biomedical literature on environmental toxicology and immunology. We introduce the Exposome Interpreter, a multimodal framework designed to infer patient-specific relationships between environmental exposures and immunological dysregulation. Our approach first employs fine-tuned Vision-Language Models (VLMs), including Gemini 2.5 Flash and PaliGemma, for high-fidelity information extraction from visually complex lab reports, canonicalizing semi-structured biomarker data into a machine-readable format. Concurrently, a Retrieval-Augmented Generation (RAG) pipeline, leveraging a domain-adapted Gemma model, queries the biomedical literature to construct a knowledge graph linking chemical agents to specific immune pathways. By integrating the structured patient data with this synthesized knowledge base and the patient's product history, the Exposome Interpreter generates ranked, evidence-backed hypotheses for environmental triggers, including direct mapping of abnormal biomarkers to specific consumer products.

1 Introduction

The prevalence of Autoimmune Diseases (ADs) is rising globally (1), yet the promise of personalized medicine remains largely unfulfilled for this population (2). Current therapeutic paradigms rely heavily on broad-spectrum immunosuppression, which manages symptoms but fails to address the underlying environmental factors that initiate and perpetuate immune dysregulation (3). The exposome concept, first proposed to capture the entirety of environmental exposures from conception onwards, offers a critical lens for understanding autoimmunity (4). This concept emphasizes that environmental factors and their corresponding biological responses are as important as genetics in determining disease risk (5). This includes diet, xenobiotics, microbial exposures, and lifestyle factors, which together provide a necessary complement to the genome for understanding chronic disease (6). However, integrating the exposome into clinical care is fundamentally an information processing challenge of immense scale. Clinicians cannot manually reconcile a patient's unique immunological biomarker profile against the vast, dynamic body of knowledge concerning thousands of environmental chemicals and their biological impacts. This challenge is characterized by two primary bottlenecks:

The clinical data standardization bottleneck: Crucial patient biomarker data and exposure history are fragmented and locked within semi-structured, visually rich documents (e.g., PDF lab reports)

2nd Workshop on Multi-modal Foundation Models and Large Language Models for Life Sciences

and unstructured inputs (e.g., product checklists or images). These inputs exhibit extreme variability in format, terminology, and layout. Traditional data ingestion methods (e.g., standard OCR or rule-based parsing) are brittle, failing to capture the necessary contextual and spatial information (e.g., associating a value with its corresponding analyte, or an ingredient with its product).

The biomedical knowledge synthesis bottleneck: The scientific literature detailing the immunotoxic effects of environmental exposures is vast, complex, and rapidly evolving. Extracting actionable, mechanistic insights requires synthesizing nuanced relationships between specific chemicals, biological pathways (e.g., Aryl Hydrocarbon Receptor activation, Th17 polarization), and clinical outcomes across diverse study designs (3).

The multi-modal reasoning capabilities of models like 2.5 Flash (7), combined with the specialized adaptability of open-source LLMs like Gemma (8), provide the necessary tools to interpret complex documents and synthesize intricate biomedical knowledge at scale. We introduce the Valence Wellbeing framework, the Exposome Interpreter, an application of these advanced Al technologies designed to decode the exposome and deliver personalized autoimmune treatment strategies.[1]

2 The Valence Wellbeing framework: A multi-modal architecture

Our framework operates through a three-stage architecture designed to (1) structure heterogeneous clinical data, (2) synthesize relevant biomedical knowledge, and (3) integrate these streams to provide personalized interventions (Figure 1).

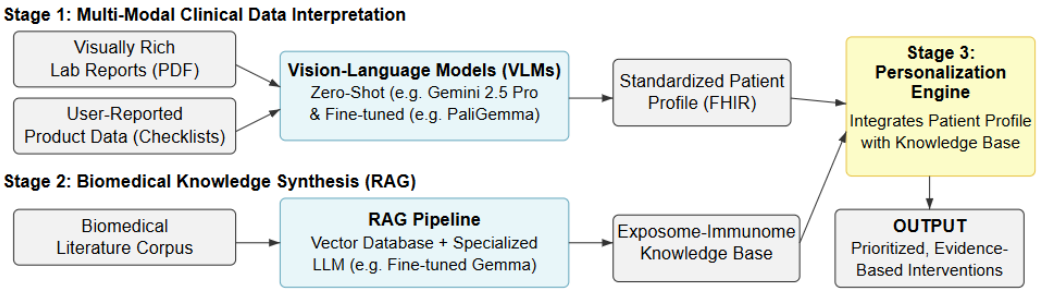

Figure 1: The exposome interpreter architecture utilizing multi-modal frameworks

2.1 Stage 1: Multi-modal clinical data interpretation

The first stage addresses the profound heterogeneity of clinical data, encompassing both clinical lab reports and patient-reported product usage. Traditional OCR methods extract raw text but fail to capture the crucial context where the interpretation of a value is dependent on its spatial relationship to headers, tabular structures, units, and reference ranges, or where an ingredient must be associated with a specific product. We apply Vision-Language Models (VLMs) to interpret these visually rich documents and inputs using a hybrid strategy of two complementary approaches:

Generalist VLMs (zero/few-shot): Leveraging the advanced capabilities of large, proprietary models such as Google's Gemini 2.5 Flash (7). These models demonstrate remarkable performance in extracting structured data directly from complex images and PDFs with minimal prompt engineering, often by specifying the desired JSON schema in the prompt.

Specialized open-source VLMs: To optimize for robustness, efficiency, and accuracy on specialized clinical terminology and rare lab formats, we are fine-tuning open-source models like PaliGemma (9).We propose using weak supervision, where annotations generated by the Generalist VLM (Gemini 2.5 Flash) are used to train the Specialized VLM (PaliGemma), accelerating the development process.

The raw data is then normalized: test names are mapped to standard ontologies like LOINC (10), and units are standardized. This process yields a machine-readable biomarker profile, formatted as FHIR-compliant data to ensure semantic interoperability (11).

[1]This paper presents the architectural framework, which is a work in progress. Links to a live implementation and relevant files, including a representative anonymized report, are available in the technical appendices.

2.2 Stage 2: Biomedical knowledge synthesis via RAG

The second stage aims to synthesize the vast literature on environmental toxicology and immunology. Traditional NLP approaches (e.g., Named Entity Recognition (NER) and Relation Extraction (RE)) often miss crucial context, pathway information, and the nuance of experimental findings (e.g., dose-dependency, experimental modality), which are essential for clinical application. We employ an Retrieval-Augmented Generation (RAG) architecture (12), allowing for comprehensive synthesis while ensuring the output is grounded in scientific evidence.

Corpus and embedding: A comprehensive corpus, including PubMed abstracts, full-text articles, and databases like the Comparative Toxicogenomics Database (CTD) [8] and EPA CompTox, is processed. We utilize specialized biomedical embedding models (e.g., those trained on PubMed literature) to generate dense vector representations of the text (13).

Advanced retrieval: When querying the impact of an exposure on a biomarker, the system performs a vector similarity search on a corpus containing resources like the Comparative Toxicogenomics Database (CTD) (14) and EPA CompTox, enhanced with a re-ranking stage to optimize the relevance of the retrieved documents.

Specialized LLM synthesis: The retrieved documents are passed to a specialized LLM (e.g., fine-tuned Gemma) to synthesize the information. This specialization enables the model to understand complex biological interactions (e.g., "Triclosan acts as an endocrine disruptor, potentially polarizing T-helper cells towards a Th17 phenotype") and generate high-fidelity summaries, including mechanisms of action and evidence strength.

This RAG system effectively constructs an "Exposome-Immunome Knowledge Base", capturing the multifaceted relationships between environmental factors and immune responses.

2.3 Stage 3: Personalization engine

The final stage integrates the standardized patient profile (Stage 1) with the synthesized knowledge base (Stage 2). The engine identifies the patient's abnormal biomarkers and queries the knowledge base for environmental factors known to influence those markers in the observed direction. A multi-factor prioritization algorithm ranks these potential triggers. The ranking considers:

Strength of evidence: Weighted by the study types synthesized by the RAG system (e.g., Randomized Controlled Trials (RCTs) > longitudinal cohorts > cross-sectional studies > in vitro studies).

Biological plausibility: Assessing whether the synthesized mechanism of action aligns with the patient's overall clinical presentation and biomarker profile.

Magnitude of effect and consistency: The effect size and reproducibility of an association in the literature. This delivers targeted, evidence-based hypotheses linking specific patient products to potential immunological dysregulation.

3 Methodology and evaluation

This section details the current methodology for the framework's development and validation. To protect patient privacy while ensuring model robustness, our development utilizes a limited set of anonymized clinical reports supplemented with synthetic data, and we provide a representative report and product image in Technical Appendices to demonstrate this approach.

3.1 Data sources

Clinical lab reports: Initial development is grounded in a small set of real-world clinical lab reports that have been fully anonymized. Recognizing that the strict privacy constraints of HIPAA limit the availability of large-scale public datasets for autoimmune conditions, we have supplemented our primary data with a publicly available collection of 426 medical reports to validate the generalizability of our VLM-based parsing engine. A representative example of our core anonymized autoimmune reports is shared with this paper to demonstrate the system's application on domain-specific data formats.

Biomedical literature corpus: A corpus compiled from the PubMed Central (PMC) open-access subset, toxicology databases (e.g., CTD, EPA CompTox), and pre-print servers.

3.2 Model training and fine-tuning

VLM fine-tuning (PaliGemma): We are fine-tuning PaliGemma using our dataset of anonymized and synthetic lab reports and product images, augmented by weak supervision from Gemini 2.5 Flash. The objective function focuses on accurate extraction of key-value pairs (biomarkers/values, ingredients/products) and successful normalization to standard ontologies.

LLM specialization (Gemma): Gemma is being fine-tuned using instruction-tuning techniques on the biomedical corpus (15). Training prompts focus on summarizing relationships between chemicals and biomarkers, emphasizing pathway information and evidence strength assessment.

3.3 Evaluation metrics

Stage 1 Evaluation: We evaluated Stage 1 key-value pair extraction on the public Kaggle Lab Report Dataset (426 diverse reports) using macro-averaged Precision and Recall. A zero-shot Gemini 2.5 Flash model achieved 0.88 Precision and 0.86 Recall. Performance improved to 0.91 Precision and 0.89 Recall with a fine-tuned PaliGemma model using weak supervision. For interoperability, the framework normalizes extracted terms to standard ontologies like LOINC and PubChem.

Stage 2 Evaluation: The RAG system is evaluated on the relevance of retrieved documents (e.g., using Mean Average Precision) and the factuality and groundedness of the synthesized responses. This ensures that the generated knowledge is both pertinent and scientifically sound.

End-to-End evaluation: The system's ability to correctly identify pre-established environment-biomarker links and map them to the correct product is evaluated using simulated patient scenarios. This comprehensive assessment validates the integration of all framework components, from data extraction to personalized hypothesis generation.

4 Challenges and future directions

Our ongoing research to bridge the exposome and immunome using LLMs is guided by several inherent challenges that define our future directions. The first challenge is compliance with the Health Insurance Portability and Accountability Act (HIPAA) for handling Protected Health Information (PHI), which legally prohibits all data sharing and analysis until our framework is validated. Once established, our focus will shift to analytical rigor, specifically establishing causality versus correlation. We plan to integrate computational causal inference techniques, such as the Bradford Hill criteria (16) and insights from Mendelian randomization studies (17), to move beyond simple associations. This analytical work is compounded by the "long-tail" problem of data heterogeneity, stemming from the sheer diversity of lab report formats. To manage this, we will employ active learning strategies and human-in-the-loop verification to efficiently address edge cases. To expand the utility of our model, we will also tackle the issue of understudied chemicals through predictive toxicology, integrating Cheminformatics Foundation Models to infer the potential immunotoxicity of novel compounds via Quantitative Structure-Activity Relationships (QSAR) (18). Finally, to ensure clinical adoption, we recognize the paramount importance of interpretability and explainability (XAI) (19) by increasing the transparency of our VLM and prioritization algorithms, complementing the RAG system's traceability to build clinical trust.

5 Conclusion

Effectively managing autoimmune conditions requires empowering individuals with a data-driven understanding of their environmental triggers. By harnessing the capabilities of advanced multimodal foundation models like Gemini 2.5 Flash and PaliGemma and the knowledge synthesis of RAG systems, we can overcome these informational challenges. The Exposome Interpreter framework demonstrates the potential of these technologies to complement conventional approaches by providing precise, evidence-based environmental health guidance, right down to the level of specific consumer products. This paves the way for a new era of proactive and personalized autoimmune wellness.

Acknowledgments and Disclosure of Funding

Use unnumbered first level headings for the acknowledgments. All acknowledgments go at the end of the paper before the list of references. Moreover, you are required to declare funding (financial activities supporting the submitted work) and competing interests (related financial activities outside the submitted work). More information about this disclosure can be found at: `https://neurips.cc/Conferences/2025/PaperInformation/FundingDisclosure`.

Do **not** include this section in the anonymized submission, only in the final paper. You can use the `ack` environment provided in the style file to automatically hide this section in the anonymized submission.

References

[1] Lerner, A., Jeremias, P., & Matthias, T. (2015). The world incidence and prevalence of autoimmune diseases is increasing. *International Journal of Celiac Disease, 3*(4), 151-155.

[2] Ginsburg, G. S., & Phillips, K. A. (2018). Precision medicine: from science to value. *Health Affairs, 37*(5), 694-701.

[3] Vojdani, A., & Pollard, K. M., Campbell, A. W. (2014). Environmental triggers and autoimmunity. *Autoimmune Diseases, 2014*.

[4] Wild, C. P. (2005). Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer epidemiology, biomarkers & prevention, 14*(8), 1847-1850.

[5] Rappaport, S. M., & Smith, M. T. (2010). Epidemiology. Environment and disease risks. *Science, 330*(6003), 460-461.

[6] Miller, G. W. (2013). *The Exposome: A Primer*. Academic Press.

[7] Team, G., et al. (2025). Gemini 2.5: Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next Generation Agentic Capabilities. *arXiv preprint arXiv:2507.06261*.

[8] Gemma Team, Google. (2024). Gemma: Open Models Based on Gemini Research and Technology. *[Technical Report]*.

[9] Chen, X., et al. (2024). PaliGemma: Open-source vision-language models. *[Technical Report]*.

[10] McDonald, C. J., Huff, S. M., Suico, J. G., Hill, G., Leavelle, D., Aller, R., ... & Mercer, K. (2003). LOINC, a universal standard for identifying laboratory and clinical observations: a 5-year update. *Clinical chemistry, 49*(4), 624-633.

[11] Bender, D., & Sartipi, K. (2013). HL7 FHIR: An agile and RESTful approach to healthcare information exchange. In *Proceedings of the 26th IEEE international symposium on computer-based medical systems* (pp. 326-331).

[12] Lewis, P., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems, 33*, 9459-9474.

[13] Lee, J., et al. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics, 36*(4), 1234-1240.

[14] Davis, A. P., et al. (2021). The Comparative Toxicogenomics Database: update 2021. *Nucleic Acids Research, 49*(D1), D1138-D1143.

[15] Wei, J., et al. (2021). Finetuned language models are zero-shot learners. *arXiv preprint arXiv:2109.01652*.

[16] Hill, A. B. (1965). The environment and disease: association or causation?. *Proceedings of the Royal Society of Medicine, 58*(5), 295.

[17] Davey Smith, G., & Hemani, G. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human molecular genetics, 23*(R1), R89-R98.

[18] Cherkasov, A., et al. (2014). QSAR modeling: where have you been? Where are you going to?. *Journal of medicinal chemistry, 57*(12), 4977-5010.

[19] Tjoa, E., & Guan, C. (2020). A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. *IEEE Transactions on Neural Networks and Learning Systems, 32*(11), 4793-4813.

A Technical Appendices and Supplementary Material

A.1 Data Availability

The Exposome Interpreter framework described in this paper is an active, ongoing project. A live implementation is available for public use at https://valencewellbeing.com. To demonstrate the framework's functionality, an anonymized lab report and a product image are attached below. The dataset of 426 diverse reports is available at https://www.kaggle.com/datasets/dikshaasinghhh/bajaj.

A.2 Licenses and Terms of Use

The assets used in our proposed framework are governed by their respective licenses and terms of use. Proprietary models like Gemini 2.5 Flash are subject to Google's Generative AI Terms of Service, while open-source models like Gemma and PaliGemma, and databases like LOINC, are used in accordance with their specific open licenses.

JANE R. DOE

JANE R. DOE

FINAL REPORT

Accession ID: 9876543210 40

Name: JANE R. DOE
JANE R. DOE
Date of Birth: 01-15-1985
Biological Sex: Female
Age: 40
Height:
Weight:
Fasting:

Telephone: 1-800-555-0199
Street Address:
123 WELLNESS WAY, ANYTOWN, CA 90210
Email: email@example.com

Provider Information

Practice Name: Community Health Clinic
Community Health Clinic
Provider Name: DR SUSAN
JONES , MD(54321)
Phlebotomist: 608

Telephone: 1-888-555-0123
Address:

Report Information

Current Result | Previous Result | In Control | Moderate | Risk

Specimen Information

Sample Type	Collection Time	Received Time	Report	Final Report Date
Metal Free Urine	2025-07-15 07:45 (PDT)	2025-07-16 13:45 (PDT)	Environmental Toxins - P2	2025-07-25 00:12 (PDT)

VibrantWellness
3521 Leonard Ct, Santa Clara, CA 95054
1-866-364-0963 | support@vibrant-america.com | www.vibrant-america.com

TNP Test not performed | R&L Refer to risks and limitations at the end of report | Notes Refer to Lab notes at the end of the table

Environmental Toxins

INTRODUCTION

Vibrant Wellness is pleased to present to you, 'Environmental Toxins Panel', to help you make healthy lifestyle, dietary and treatment choices in consultation with your healthcare provider. It is intended to be used as a tool to encourage a general state of health and well-being.

The Vibrant Environmental Toxins Panel is a test to measure levels of Environmental Toxins that someone might be exposed to. The panel is designed to give a complete picture of an individual's levels of these toxins in urine. The panel is sub-grouped into Pesticides, Phthalates, Parabens, Acrylic, Alkyl phenols and Volatile Organic Compounds. Reference ranges for tests flagged with ^ were determined based on NHANES data (cdc.gov/nhanes) if available and other reference ranges are established based on urine samples from 1000 apparently healthy individuals.

Methodology:

The Vibrant Environmental Toxins panel uses tandem mass spectrometry methodology (LC-MS/MS) for quantitative detection of toxins in urine samples. Urine creatinine is measured using a kinetic colorimetric assay based on the Jaffé method. All environmental toxins are reported as the quantitative result normalized to urine creatinine to account for urine dilution variations.

Interpretation of Report:

The report begins with the summary page which lists only the environmental toxins whose levels are high or moderate in the reference range. Additionally, the previous value is also indicated to help check for improvements every time the test is ordered. Following this section is the complete list of the environmental toxins and their absolute levels are normalized with respect to Creatinine in a histogram format to enable a full overview along with the reference ranges. The level of the environmental toxins is shown with three shades of color – Green, Yellow and Red. The result in green corresponds to 0th to 75th percentile indicates mild exposure to the respective toxin. The result in yellow corresponds to 75th to 95th percentile indicates moderate exposure to the respective toxin whereas the result in red corresponding to greater than 95th percentile indicates high exposure to the respective toxin. All contents provided in the report are purely for informational purposes only and should not be considered medical advice. Any changes based on the information should be made in consultation with the clinical provider.

The Vibrant Wellness platform provides tools for you to track and analyze your general wellness profile. Testing for the Environmental Toxins panel is performed by Vibrant America, a CLIA certified lab CLIA#:05D2078809. Vibrant Wellness provides and makes available this report and any related services pursuant to the Terms of Use Agreement (the "Terms") on its website at www.vibrant-wellness.com. By accessing, browsing, or otherwise using the report or website or any services, you acknowledge that you have read, understood, and agree to be bound by these terms. If you do not agree to accept these terms, you shall not access, browse, or use the report or website. The statements in this report have not been evaluated by the Food and Drug Administration and are only meant to be lifestyle choices for potential risk mitigation. Please consult your healthcare provider for medication, treatment, or lifestyle management. This product is not intended to diagnose, treat, or cure any disease.

Please note:

Pediatric ranges have not been established for this test. It is important that you discuss any modifications to your diet, exercise, and nutritional supplementation with your healthcare provider before making any changes.

Environmental Toxins - Summary

Environmental phenols

Test Name	Current	Previous	Result (75th / 95th)	Reference
4-Nonylphenol (ug/g)	2.86	3.76 (06-15-2024)	0.42 / 2.06	≤2.06

BACKGROUND

4-Nonylphenols are used in manufacturing antioxidants, lubricating oil additives, laundry and dish detergents, emulsifiers, and solubilizers. These compounds are also precursors used to produce paints, pesticides, cosmetics, and plastics. Nonylphenol persists in aquatic environments and is moderately bio accumulative. It is not readily biodegradable, and it can take months or longer to degrade in surface waters, soils, and sediments.

ASSOCIATED RISK

It has a potential role as an endocrine disruptor and xenoestrogen due to its ability to act with estrogen-like activity. Nonylphenol exposure has also been associated with breast cancer. Exposure to 4-nonylphenol, is known to cause some long-term behavioural abnormalities, including autism spectrum disorder.

POSSIBLE SOURCES

Human exposure to 4-nonylphenol primarily occurs through the consumption of contaminated food and water. 4-nonylphenol are notably found in fish and shellfish, as these aquatic organisms can absorb them from their environment. Additionally, when sewage sludge is used as fertilizer on agricultural land, 4-nonylphenol can be introduced into the soil, potentially affecting crops and livestock.

DETOX SUGGESTIONS

To detoxify 4-Nonylphenols from the body, focus on consuming foods rich in antioxidants, such as fruits and vegetables, which aid in neutralizing and eliminating toxins. Hydration is also crucial to support the body's natural detox processes, so drink plenty of water to flush out toxins through urine and sweat.

Test Name	Current	Previous	Result (75th / 95th)	Reference
Triclosan (TCS)^ (ug/g)	40.85	29.91 (06-15-2024)	29.9 / 358	≤358

BACKGROUND

Triclosan (TCS) is an antibacterial and antifungal agent present in some consumer products, including toothpaste, soaps, detergents, toys, and surgical cleaning treatments.

ASSOCIATED RISK

TCS has been linked to an increased risk of food allergies, adding to concerns about its potential health effects. Furthermore, TCS has been identified as a weak endocrine disruptor, suggesting its ability to interfere with hormonal balance. Notably, prenatal exposure to triclosan has been associated with elevated cord testosterone levels in infants, highlighting its potential impact on early development and hormonal regulation. Exposure to this toxin has been linked to early kidney injury, an elevated risk of chronic kidney disease (CKD), and the potential for end-stage renal disease (ESRD). It is also responsible for inducing hepatic toxicity, renal toxicity, intestinal damage, and impairment of thyroid function.

POSSIBLE SOURCES

Exposure to triclosan occurs through skin absorption during activities like handwashing and showering, as well as through ingestion via tooth brushing, mouthwash, and swallowing, with additional potential sources including consuming plants grown in sewage sludge-treated soil and fish exposed to triclosan.

DETOX SUGGESTIONS

Incorporating binders like charcoal or clay-based products aids in reducing toxin levels by effectively binding and eliminating environmental toxins from the body. These substances encapsulate toxins, such as heavy metals and pollutants, facilitating their removal and potentially reducing zonulin levels, which contribute to a leaky gut (16). Supplementing with antioxidants like glutathione is essential for protecting cells from oxidative damage induced by environmental toxins. Glutathione, the body's primary antioxidant and detoxifier, plays a crucial role in combating harmful free radicals, supporting Phase II detoxification pathways, and preventing deficiency-related health issues.

Environmental Toxins - Summary

Herbicides

Test Name	Current	Previous	Result (75th / 95th)	Reference
Glyphosate (ug/g)	4.98	13.12 (06-15-2024)	1.65 / 7.6	≤7.6

BACKGROUND

Glyphosate is a broad-spectrum systemic herbicide and crop desiccant widely utilized to eliminate weeds, particularly annual broadleaf weeds and competing grasses in crop fields.

ASSOCIATED RISK

This exposure may have implications for liver health, metabolic disorders, and adverse effects on the nervous system. Glyphosate exposure during early life stages can disrupt normal cell development, impacting critical signalling pathways and causing issues like altered differentiation, neuronal growth, migration, and myelination (2,3).

POSSIBLE SOURCES

Glyphosate exposure can stem from various sources, including occupational use, residential proximity to farmland, living with occupational users, dietary consumption of food with residues, ingesting contaminated water, and secondary exposure through contact with treated areas.

DETOX SUGGESTIONS

Citrus pectin, alginates from kelp, and glycine act as powerful detoxifiers. Citrus pectin clears environmental toxins and cholesterol, alginates protect against herbicides and remove toxins, while glycine aids in glutathione production, preventing glyphosate storage. Gingko biloba serves as a potent protector against glyphosate toxicity (20-22).

Mitochondrial Marker

Test Name	Current	Previous	Result (75th / 95th)	Reference
Tiglylglycine (TG) (ug/g)	0.10	0.12 (06-15-2024)	0.09 / 3.24	≤3.24

BACKGROUND

Tiglylglycine (TG) is associated with both mitochondrial and/or genetic disorders. It is a specific metabolite that plays a crucial role in the diagnosis of a rare genetic disorder known as '3-Hydroxyisobutyryl-CoA Hydrolase (HIBCH) Deficiency.' HIBCH deficiency is an inborn error in isoleucine metabolism, leading to the accumulation of isoleucine metabolites, including TG, in the urine of affected individuals

ASSOCIATED RISK

Mutations of mitochondrial DNA can be triggered by toxins, infections, inflammation, and nutritional deficiencies. Mitochondrial dysfunction has been linked with aging, diabetes, autism, chronic fatigue syndrome, PD and Alzheimer's syndromes. The presence of elevated levels of TG in the urine serves as a biomarker for HIBCH deficiency. This disorder is associated with various clinical manifestations, including microcephaly, epilepsy, choreoathetoid movements, ophthalmologic disorders, progressive neurodegeneration, psychomotor retardation or regression, hearing impairment, and even cardiomyopathy. Unfortunately, the disease can lead to a significantly shortened lifespan for some individuals

POSSIBLE SOURCES

β-ketothiolase deficiency is a rare genetic disorder characterized by the inability to properly metabolize certain compounds, including isoleucine and its derivatives. Therefore, individuals with β -ketothiolase deficiency usually excrete TG in excess amounts.

DETOX SUGGESTIONS

Tiglylglycine (TG) can be detoxified from the body through enzymatic breakdown pathways in the liver, where it is metabolized into smaller molecules that can be excreted through urine. Adequate hydration and a balanced diet rich in nutrients that support liver function can aid in the efficient removal of TG from the body.

Environmental Toxins - Summary

Other Markers

No markers are outside the normal reference range

Parabens

No markers are outside the normal reference range

Pesticides

Test Name	Current	Previous	Result (75th / 95th)	Reference
2,2-bis(4-Chlorophenyl) acetic acid (DDA) (ug/g)	8.88	13.57 (06-15-2024)	7.9 / 19	≤19

BACKGROUND

DDT metabolism in humans yields 2,2-bis (4-chlorophenyl) acetic acid (DDA) as the principal urinary metabolite and potential exposure biomarker. DDT is a persistent organic pollutant that is readily adsorbed to soils and sediments, which can act both as sinks and as long-term sources of exposure. DDT was a commonly used pesticide for insect control. DDT was used to control malaria and typhus.

ASSOCIATED RISK

DDT is an endocrine disruptor and indicates possible disruption in semen quality, menstruation, gestational length, and duration of lactation. Chronic exposure to DDT will build up in areas of the body with high lipid content and can affect reproductive capabilities and the embryo or fetus. It is considered likely to be a human carcinogen, especially for breast cancer. DDE is a metabolite of DDT and is excreted as DDA in the urine

POSSIBLE SOURCES

DDT can be absorbed by humans through inhalation of gaseous and particulate phases, direct dermal contact, ingestion of contaminated substances, and exposure to contaminated soil or products.

DETOX SUGGESTIONS

DDT can accumulate in the body and have been associated with adverse health effects. Sweating induced by infrared and steam sauna sessions can help eliminate pesticides from the body. As with other toxins, sweating allows pesticides to be excreted through the skin.

Phthalates

No markers are outside the normal reference range

Volatile organic compounds

No markers are outside the normal reference range

Creatinine

Test Name	Current	Previous	Result	Reference
Urine Creatinine (mg/mL)	0.76	1.37 (06-15-2024)	0 / 0.24 / 2.16	0.25-2.16

Patient Name: JANE R. DOE
Date of Birth: 01-15-1985 Accession ID: 9876543210 40
Service Date: 2025-07-15 07:45 (PDT)

Environmental Toxins

Environmental phenols

Test Name	Current	Previous	Result (75th / 95th)	Reference
4-Nonylphenol (ug/g)	2.86	3.76 (06-15-2024)	0.42 / 2.06	≤2.06
Bisphenol A (BPA)^ (ug/g)	1.51	1.73 (06-15-2024)	2.12 / 5.09	≤5.09
Triclosan (TCS)^ (ug/g)	40.85	29.91 (06-15-2024)	29.9 / 358	≤358

Herbicides

Test Name	Current	Previous	Result (75th / 95th)	Reference
2,4-Dichlorophenoxyacetic Acid (2,4-D)^ (ug/g)	0.30	0.02 (06-15-2024)	0.5 / 1.55	≤1.55
Atrazine ^ (ug/g)	<0.01	<0.01 (06-15-2024)	0.02 / 0.05	≤0.05
Atrazine mercapturate^ (ug/g)	<0.01	0.02 (06-15-2024)	0.02 / 0.05	≤0.05
Glyphosate (ug/g)	4.98	13.12 (06-15-2024)	1.65 / 7.6	≤7.6

Mitochondrial Marker

Test Name	Current	Previous	Result (75th / 95th)	Reference
Tiglylglycine (TG) (ug/g)	0.10	0.12 (06-15-2024)	0.09 / 3.24	≤3.24

Other Markers

Test Name	Current	Previous	Result (75th / 95th)	Reference
Diphenyl Phosphate (DPP) (ug/g)	0.26	0.74 (06-15-2024)	1.1 / 3.7	≤3.7
N-acetyl-S-(2-carbamoylethyl)-cysteine^ (ug/g)	40.03	15.17 (06-15-2024)	82 / 199	≤199
Perchlorate (PERC)^ (ug/g)	2.30	0.31 (06-15-2024)	4.89 / 10.7	≤10.7

Parabens

Test Name	Current	Previous	Result (75th / 95th)	Reference
Butylparaben^ (ug/g)	0.20	0.18 (06-15-2024)	0.25 / 4.39	≤4.39
Ethylparaben ^ (ug/g)	0.10	2.14 (06-15-2024)	5.40 / 99.3	≤99.3
Methylparaben^ (ug/g)	6.00	73.23 (06-15-2024)	180 / 653	≤653
Propylparaben^ (ug/g)	2.04	16.08 (06-15-2024)	36.7 / 222	≤222

Environmental Toxins

Pesticides

Test Name	Current	Previous	Result (75th / 95th)	Reference
2,2-bis(4-Chlorophenyl) acetic acid (DDA) (ug/g)	8.88	13.57 (06-15-2024)	7.9 / 19	≤19
3-Phenoxybenzoic Acid (3PBA)^ (ug/g)	0.77	0.88 (06-15-2024)	1.01 / 5.44	≤5.44
Diethyl phosphate (DEP)^ (ug/g)	0.27	2.71 (06-15-2024)	3.2 / 15.7	≤15.7
Diethyldithiophosphate (DEDTP)^ (ug/g)	0.11	0.03 (06-15-2024)	0.17 / 0.3	≤0.3
Diethylthiophosphate (DETP)^ (ug/g)	0.99	0.07 (06-15-2024)	1.24 / 3.92	≤3.92
Dimethyl phosphate (DMP)^ (ug/g)	5.36	8.50 (06-15-2024)	9.1 / 33.6	≤33.6
Dimethyldithiophosphate (DMDTP)^ (ug/g)	0.64	0.10 (06-15-2024)	0.67 / 6.12	≤6.12
Dimethylthiophosphate (DMTP)^ (ug/g)	1.20	20.55 (06-15-2024)	5.91 / 33.7	≤33.7

Phthalates

Test Name	Current	Previous	Result (75th / 95th)	Reference
Mono-(2-ethyl-5-hydroxyhexyl) phthalate (MEHHP)^ (ug/g)	3.26	6.70 (06-15-2024)	14.1 / 37.7	≤37.7
Mono-(2-ethyl-5-oxohexyl) phthalate (MEOHP)^ (ug/g)	7.68	4.62 (06-15-2024)	8.99 / 23.4	≤23.4
Mono-2-ethylhexyl phthalate (MEHP)^ (ug/g)	0.38	2.22 (06-15-2024)	2.73 / 8.47	≤8.47
Mono-ethyl phthalate (MEtP)^ (ug/g)	4.34	0.09 (06-15-2024)	94.2 / 540	≤540

Volatile organic compounds

Test Name	Current	Previous	Result (75th / 95th)	Reference
2-Hydroxyethyl Mercapturic Acid (HEMA)^ (ug/g)	0.13	0.91 (06-15-2024)	1.7 / 4.75	≤4.75
2-Hydroxyisobutyric Acid (2HIB) (ug/g)	577.90	335.44 (06-15-2024)	795.93 / 1215.72	≤1215.72
2-Methylhippuric Acid (2MHA)^ (ug/g)	57.07	0.04 (06-15-2024)	77.9 / 248	≤248
3-Methylhippuric Acid (3MHA) (ug/g)	45.59	0.30 (06-15-2024)	64.8 / 612.83	≤612.83
4-Methylhippuric Acid (4MHA) (ug/g)	56.59	0.20 (06-15-2024)	65.51 / 752.72	≤752.72
N-Acetyl (2-Cyanoethyl) Cysteine (NACE)^ (ug/g)	3.49	0.04 (06-15-2024)	5.28 / 256	≤256
N-Acetyl (2,Hydroxypropyl) Cysteine (NAHP)^ (ug/g)	61.85	58.87 (06-15-2024)	101 / 403	≤403

Environmental Toxins

Volatile organic compounds

Test Name	Current	Previous	Result (75th / 95th)	Reference
N-Acetyl (3,4-Dihydroxybutyl) Cysteine^ (ug/g)	125.68	0.24 (06-15-2024)	374 / 583	≤583
N-Acetyl (Propyl) Cysteine (NAPR)^ (ug/g)	7.55	0.03 (06-15-2024)	11.3 / 46.1	≤46.1
N-acetyl phenyl cysteine (NAP)^ (ug/g)	0.74	1.24 (06-15-2024)	1.29 / 3.03	≤3.03
Phenyl glyoxylic Acid (PGO)^ (ug/g)	156.01	137.08 (06-15-2024)	285 / 518	≤518

Risk and Limitations

This test has been developed and its performance characteristics determined by Vibrant America LLC., a CLIA certified lab. These assays have not been cleared or approved by the U.S. Food and Drug Administration.

Vibrant Environmental Toxins panel does not demonstrate absolute positive and negative predictive values for any condition. Its clinical utility has not been fully established. Clinical history and current symptoms of the individual must be considered by the healthcare provider prior to any interventions. Test results should be used as one component of a physician's clinical assessment.

Environmental Toxins Panel testing is performed at Vibrant America, a CLIA certified laboratory and utilizes ISO-13485 developed technology. Vibrant America has effective procedures in place to protect against technical and operational problems. However, such problems may still occur. Examples include failure to obtain the result for a specific toxin due to circumstances beyond Vibrant's control. Vibrant may re-test a sample in order to obtain these results but upon re-testing the results may still not be obtained. As with all medical laboratory testing, there is a small chance that the laboratory could report incorrect results. A tested individual may wish to pursue further testing to verify any results.

The information in this report is intended for educational purposes only. While every attempt has been made to provide current and accurate information, neither the author nor the publisher can be held accountable for any errors or omissions.

Vibrant Wellness makes no claims as to the diagnostic or therapeutic use of its tests or other informational materials. Vibrant Wellness reports and other information do not constitute the giving of medical advice and are not a substitute for a professional healthcare practitioner. Please consult your provider for questions regarding test results, or before beginning any course of medication, supplementation or dietary/lifestyle changes. Users should not disregard, or delay in obtaining, medical advice for any medical condition they may have, and should seek the assistance of their health care professionals for any such conditions.

Example Product Image: The figure below shows an example of a consumer product image that a user might provide.

Figure 2: Example of a user-submitted product image.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction claim to introduce the Exposome Interpreter, a multi-modal framework using Vision-Language Models (VLMs) and Retrieval-Augmented Generation (RAG) for personalized autoimmune care. This accurately reflects the architecture detailed in Section 2 and the methodology in Section 3, which form the core contribution of the paper.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Section 4, titled "Challenges and future directions," explicitly discusses several limitations and challenges of the proposed framework, including the difficulty of obtaining the compliance, establishing causality vs. correlation, data heterogeneity, and the need for interpretability (XAI).

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

   Justification: The paper proposes a system architecture and application framework rather than presenting novel theoretical results; therefore, it does not include theorems or mathematical proofs.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [NA]

   Justification: The paper describes the complete framework and does not include any experimental results.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: This paper proposes a novel framework and does not contain formal experiments or numerical results to reproduce. However, to provide concrete examples of the data modalities and the real-world context for our work, we have included all the information and the links. These include: A link to the Valence Wellbeing website (https://www.valencewellbeing.com) to offer context on the project's application and motivation. An anonymized, real-world example of a clinical lab report and link to 426 reports publically available datasets (https://www.kaggle.com/datasets/dikshaasinghhh/bajaj) that the proposed VLM stage is designed to interpret. An example of a consumer product image, representing the type of user-provided data the system is built to handle..

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [NA]

   Justification: The paper does not include experimental results. While Section 3 outlines the methodology for VLM fine-tuning and LLM specialization, specific implementation details like hyperparameters are not available as the experiments have not been conducted.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [NA]

   Justification: The paper does not include any experiments or quantitative results.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
   - If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [NA]

   Justification: The paper does not include any experiments.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The proposed research aims to improve personalized healthcare for autoimmune diseases. The framework design presented in this paper conforms to the Code of Ethics.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer:

    Justification: While this paper focuses on the positive societal impact of empowering individuals with autoimmune conditions, it also addresses potential negative impacts through planned mitigation strategies. Key risks include fairness and bias in data, which will be managed via human-in-the-loop verification; significant privacy and security concerns, addressed by a foundational commitment to HIPAA compliance; and the potential for misinformation. The latter will be mitigated by ensuring the RAG system provides traceable citations and by positioning the tool as an informational wellness resource designed to support not replace the guidance of a qualified healthcare professional..

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.

- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [Yes]

    Justification: While no models or data are released with this paper, the framework is designed with safeguards for its eventual responsible deployment. The planned approach includes: (1) deploying the model via controlled access on a secure platform, not open release; (2) strict data governance within a planned HIPAA-compliant environment; and (3) inherent model safeguards, such as the proposed RAG-based traceability and human-in-the-loop verification, to ensure accountability and mitigate misuse.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The paper properly credits all assets used (e.g., Gemini, Gemma, LOINC). A statement on their governing licenses and terms of use is provided in the Technical Appendices.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: This paper describes a work in progress and does not include the release of any new assets (datasets, code, or models).

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.

- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The research presented in this paper does not involve crowdsourcing or experiments with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer: [Yes]

    Justification: The usage of LLMs (e.g., Gemma) and VLMs (e.g., 2.5 Flash, PaliGemma) is central to the proposed methodology. Section 2 details their roles in clinical data interpretation (Stage 1) and biomedical knowledge synthesis via RAG (Stage 2).

    Guidelines:

    - The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.