Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Explainable Artificial Intelligence for Predictive Toxicology and Public Health Risk Assessment: A Data-Driven Framework for Early Detection and Decision Support

Nnaemeka Kingsley Ugwumba

nnaemeka_ugwumba@uniport.edu.ng

Laskenta Technologies Limited https://orcid.org/0009-0000-2493-9846

Peter Sunday Jaja

peter_jaja@uniport.edu.ng

Laskenta Technologies Limited

Juan Sebastian Murillejo Contreras

sebastianmurillejocontreras@gmail.com

Laskenta Technologies Limited

Kelechi Ernest Okechukwu

okechukwukelechi042@gmail.com

Laskenta Technologies Limited

Research Article

Keywords: Explainable models, predictive toxicology, public health risk assessment, machine learning in health, environmental exposure analysis, interpretable systems, decision support systems, computational toxicology, health data integration, early risk detection

Posted Date: March 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9226177/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Explainable Artificial Intelligence for Predictive Toxicology and Public Health Risk Assessment: A Data-Driven Framework for Early Detection and Decision Support

AUTHOR & AFFILIATED INSTITUTION

1.Nnaemeka Kingsley Ugwumba

Institution: Laskenta Technologies Limited, Nigeria

Email Address: support@laskentatech.pro

Orcid ID: https://orcid.org/0009-0000-2493-9846

2. Peter Sunday Jaja

Email Address: peter_jaja@uniport.edu.ng

Institution: Laskenta Technologies Limited, Nigeria

Corresponding author: Nnaemeka Kingsley Ugwumba

3. Juan Sebastian Murillejo Contreras

Email: sebastianmurillejocontreras@gmail.com

Institution: Laskenta Technologies Limited, Colombia

4. Kelechi Ernest Okechukwu

Email: okechukwukelechi042@gmail.com

Institution: Laskenta Technologies Limited, Nigeria

Abstract

Public health systems continue to face increasing challenges from environmental and chemical exposures that contribute to disease burden and population level risks. Traditional toxicological assessment methods are often limited by high cost, long experimental timelines, and difficulties in translating laboratory findings into real world decision making. This study presents a data driven framework that integrates explainable artificial intelligence techniques into predictive toxicology for improved public health risk assessment. The proposed approach combines machine learning models with interpretable mechanisms to support early detection of toxicological risks while maintaining transparency in model predictions. Multi source datasets including

environmental exposure records, clinical health data, and chemical toxicity profiles are utilized to develop and validate the framework. The study demonstrates how interpretable predictive models can enhance risk classification accuracy and support evidence based public health interventions. Findings suggest that integrating explainability into predictive systems improves trust, usability, and policy relevance in toxicological applications. The framework contributes to advancing computational toxicology and offers practical implications for health agencies, researchers, and decision makers seeking timely and reliable risk assessment tools.

Keywords

Explainable models, predictive toxicology, public health risk assessment, machine learning in health, environmental exposure analysis, interpretable systems, decision support systems, computational toxicology, health data integration, early risk detection

1. Introduction

Public health and toxicology are closely linked disciplines concerned with understanding and mitigating the effects of harmful exposures on human populations. Rapid industrialization, urban expansion, and increased chemical usage have led to a rise in exposure to potentially hazardous substances. These exposures contribute to a wide range of health outcomes, including respiratory diseases, cancers, neurological disorders, and endocrine disruptions. Effective risk assessment is therefore essential for protecting population health and guiding regulatory decisions.

Traditional toxicological methods rely heavily on laboratory experiments, animal testing, and epidemiological studies. While these approaches have contributed significantly to scientific knowledge, they present several limitations. Experimental procedures are often time consuming and resource intensive, and translating findings from controlled environments to real world scenarios remains challenging. In addition, the growing volume of environmental and health related data has outpaced the capacity of conventional analytical methods to extract meaningful insights.

Recent advances in computational methods have introduced new opportunities for addressing these challenges. Machine learning techniques have shown promise in identifying patterns within large datasets

and generating predictive models for health outcomes. In toxicology, these methods have been applied to predict chemical toxicity, assess exposure risks, and support regulatory decision making. However, many predictive models operate as opaque systems, making it difficult for researchers and policymakers to understand how predictions are generated.

The lack of interpretability in predictive systems presents a critical barrier to their adoption in public health contexts. Decision makers require transparent and explainable evidence to justify interventions, develop policies, and communicate risks to stakeholders. Without clear explanations, even highly accurate models may be viewed with skepticism and remain underutilized.

This study addresses this gap by proposing a data driven framework that integrates explainable modeling techniques into predictive toxicology for public health risk assessment. The aim is to enhance both the accuracy and interpretability of predictive systems, enabling more effective early detection of toxicological risks. By combining diverse data sources and incorporating interpretable mechanisms, the framework seeks to provide actionable insights that support informed decision making.

The contributions of this study are threefold. First, it presents a structured approach for integrating multiple data sources relevant to toxicology and public health. Second, it develops predictive models that prioritize both performance and interpretability. Third, it demonstrates how explainable outputs can support practical decision making in risk assessment scenarios.

The remainder of this paper is organized as follows. Section two reviews related work in predictive toxicology and computational methods in public health. Section three describes the methodology and data sources used in the study. Section four presents the experimental results and evaluation. Section five discusses the implications of the findings. Section six concludes the study and outlines directions for future research.

2. Literature Review

The integration of computational methods into toxicology and public health has evolved significantly over the past decade, driven by the increasing availability of large-scale health and environmental datasets. Predictive modeling approaches are now

widely explored as alternatives to traditional experimental methods, particularly in addressing limitations related to cost, scalability, and translational relevance. Recent studies emphasize that conventional toxicological assessments often fail to accurately predict human specific outcomes due to interspecies differences and limited data integration capabilities (Ajisafe et al., 2025). This has created a strong motivation for the adoption of data driven approaches that can process complex and heterogeneous information sources.

2.1 Predictive Modeling in Toxicology

Predictive toxicology has emerged as a critical domain within computational health sciences, focusing on the estimation of adverse biological effects of chemical exposures using statistical and computational techniques. Modern approaches rely on diverse data sources, including chemical structure descriptors, biological assay data, and clinical health records. Large scale initiatives such as the United States Environmental Protection Agency ToxCast program have significantly advanced the field by providing high throughput screening data for thousands of chemicals, enabling the development of robust predictive models (Zhang et al., 2025).

Recent research demonstrates that machine learning models can effectively identify toxicity patterns and predict adverse outcomes at earlier stages compared to traditional laboratory-based approaches. These models leverage complex feature representations and nonlinear relationships to uncover hidden associations between exposures and health outcomes. Furthermore, predictive systems have been shown to reduce reliance on animal testing while improving efficiency in chemical risk assessment processes (Luechtefeld & Hartung, 2025).

Despite these advancements, several challenges persist. Data quality, missing values, and inconsistencies across datasets remain significant barriers to model reliability. Additionally, many predictive models are optimized primarily for performance, often at the expense of interpretability, which limits their applicability in regulatory and public health contexts.

2.2 Explainability and Interpretability in Health-Related Models

The need for transparency in predictive systems has led to growing interest in explainability within computational health research. Interpretability refers to the ability to understand and explain how a model generates its predictions, which is particularly important in high stakes domains such as healthcare and public policy. Studies have shown that the absence of clear explanations reduces trust among stakeholders and hinders the adoption of predictive systems in real world settings (Alkhanbouli et al., 2025).

Recent developments have introduced model agnostic and model specific interpretation techniques that provide insights into feature contributions and decision pathways. Methods such as local explanation techniques and feature attribution strategies enable researchers to identify key variables influencing predictions. In toxicology, these approaches have been applied to reveal how demographic, environmental, and biological factors contribute to exposure related risks.

For example, a recent study on lead toxicity prediction demonstrated how interpretable modeling techniques can enhance understanding of exposure risks in vulnerable populations. The study applied a classification model to predict lead exposure using demographic data and incorporated explanation methods to identify the most influential factors, achieving high predictive accuracy while maintaining transparency (Chaurasia et al., 2025). Such findings highlight the importance of combining predictive performance with interpretability to support clinical and public health decision making.

2.3 Applications in Public Health Risk Assessment

Public health risk assessment involves evaluating the likelihood and severity of adverse health outcomes resulting from environmental, biological, or chemical exposures. The increasing complexity of health data has necessitated the use of computational tools capable of integrating multiple data sources and generating actionable insights.

Recent studies indicate that predictive modeling approaches can support early

detection of health risks, enabling timely interventions and improved resource allocation. In community health settings, interpretable predictive systems have been used to assess frailty risks and guide preventive strategies, demonstrating the practical value of combining predictive modeling with explainability (Kim et al., 2025).

In toxicology, predictive systems are increasingly used to evaluate chemical safety, identify high risk populations, and inform regulatory decisions. These systems are particularly valuable in scenarios where rapid assessment is required, such as environmental exposure monitoring and outbreak response. The integration of multiple data modalities, including environmental data, clinical records, and sociodemographic information, has further enhanced the accuracy and relevance of these models.

2.4 Emerging Trends and Research Gaps

Recent literature highlights several emerging trends shaping the future of predictive toxicology and public health research. One key trend is the shift toward multimodal data integration, where models combine heterogeneous data sources to improve predictive accuracy and robustness. Another important development is the emphasis on human relevant models that better reflect real world biological processes, thereby improving the reliability of predictions (Luechtefeld & Hartung, 2025).

In addition, there is growing recognition of the need for standardized validation frameworks to ensure the reliability and generalizability of predictive models. The development of evaluation principles that emphasize transparency, reproducibility, and applicability is considered essential for advancing the field.

However, significant gaps remain. Many existing studies focus on model development without adequately addressing issues related to deployment, usability, and policy integration. Furthermore, there is limited research on how interpretable predictive systems can be systematically incorporated into public health decision making processes.

2.5 Summary of Literature

The reviewed literature demonstrates that predictive modeling approaches have substantial potential to transform toxicology

and public health risk assessment. Advances in data integration and computational techniques have enabled the development of models capable of identifying complex exposure related risks. At the same time, the importance of interpretability has become increasingly evident, as transparent models are more likely to be adopted in real world applications.

Despite these advancements, there remains a need for frameworks that effectively combine predictive accuracy with interpretability while addressing practical challenges related to data quality, validation, and implementation. This study seeks to address these gaps by proposing a structured approach for integrating interpretable predictive models into public health toxicology.

3. Methodology

This section presents the formal structure of the proposed framework, including data representation, model formulation, training procedures, and evaluation strategy. The methodology is designed to support predictive toxicology and public health risk assessment through a transparent and analytically rigorous approach.

3.1 Study Design and Problem Formulation

Let the dataset be defined as:

$$D = \{(x_i, y_i)\}_{i=1}^{N}$$

where $x_i \in \mathbb{R}^d$ represents the feature vector associated with the i-th sample, and $y_i \in \{0,1\}$ denotes the corresponding class label indicating the presence or absence of toxicological risk.

The objective is to learn a predictive function:

$$f: \mathbb{R}^d \to [0,1]$$

such that:

$$\hat{y}_i = f(x_i)$$

where $\hat{y}_i$ represents the estimated probability of toxicological risk.

The model is trained to minimize a loss function $\mathcal{L}$, defined as the binary cross entropy:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(f(x_i)) + (1 - y_i)\log(1 - f(x_i))]$$

3.2 Data Sources and Feature Construction

The framework integrates heterogeneous data sources relevant to toxicology and public health, including:

1. Chemical descriptors derived from molecular structure data
2. Environmental exposure records such as air and water quality indices
3. Clinical and demographic variables from population health datasets

Each sample vector x_i is constructed as:

$$x_i = [c_i, e_i, h_i]$$

where:

- c_i represents chemical features
- e_i represents environmental exposure variables
- h_i represents human health and demographic attributes

To ensure numerical stability and comparability, all continuous features are standardized:

$$x_i' = \frac{x_i - \mu}{\sigma}$$

where μ and σ denote the mean and standard deviation computed over the training set.

Missing values are handled using imputation techniques based on feature distributions, while categorical variables are encoded using appropriate transformation schemes.

3.3 Model Architecture

The overall structure of the proposed framework is illustrated in Figure 1. The framework integrates heterogeneous data sources, including chemical, environmental, and population health data, into a unified processing pipeline. The system performs data preprocessing and normalization before passing the inputs through the predictive model. In addition to generating risk predictions, the framework includes an explanation module that provides feature level insights into model outputs.

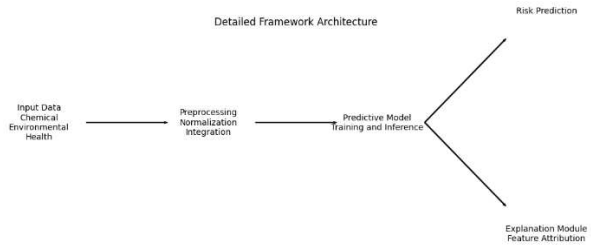


Figure 5. Detailed framework illustrating data integration, predictive modeling, and interpretability components.

As shown in Figure 1, the separation between predictive and explanatory components ensures that model outputs remain both accurate and interpretable.

The predictive function $f(x)$is modeled using a composite structure consisting of a base learner and an interpretability layer.

3.3.1 Base Predictive Model

The base model is defined as:

$$f(x) = \sigma(W^{(L)} \cdot g^{(L-1)}(\dots g^{(1)}(x)))$$

where:

- $g^{(k)}$denotes intermediate transformations
- $W^{(L)}$represents the final weight matrix
- $\sigma$is the sigmoid activation function

For structured tabular data, gradient boosted decision functions or layered nonlinear mappings are employed to capture complex interactions among variables.

3.4 Interpretability Mechanism

To ensure transparency, the framework incorporates a feature attribution function:

$$\phi_j(x) = \frac{\partial f(x)}{\partial x_j}$$

where $\phi_j(x)$quantifies the contribution of feature $x_j$to the prediction.

Additionally, local approximation techniques are used to construct interpretable surrogate models:

$$f(x) \approx g(x) = \beta_0 + \sum_{j=1}^{d} \beta_j \, x_j$$

where $g(x)$provides a linear approximation of the model in the neighborhood of a given instance.

These explanations allow identification of key risk factors influencing predictions, thereby supporting decision making in public health contexts.

3.5 Training Procedure

The model parameters are optimized using gradient based methods:

$$\theta^* = \arg\min_{\theta} \mathcal{L}(D; \theta)$$

Regularization is applied to prevent overfitting:

$$\mathcal{L}_{reg} = \mathcal{L} + \lambda \,||\, \theta \,||^2$$

where λcontrols the strength of regularization.

The dataset is partitioned into training, validation, and test subsets:

$$D = D_{train} \cup D_{val} \cup D_{test}$$

Cross validation techniques are employed to ensure robustness and generalizability of the model.

3.6 Evaluation Metrics

Model performance is assessed using multiple quantitative measures:

- Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- Recall:

$$\text{Recall} = \frac{TP}{TP + FN}$$

- F1 Score:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

- Area under the receiver operating characteristic curve:

$$AUC = \int_0^1 TPR(FP\, R^{-1}(t))dt$$

In addition to predictive performance, interpretability is evaluated based on consistency and stability of feature attribution scores across different samples.

3.7 Implementation Considerations

The framework is implemented using standard computational libraries for data processing and model training. Care is taken to ensure reproducibility through controlled initialization, consistent data splits, and parameter tuning strategies.

Computational complexity is managed by optimizing feature selection and reducing dimensionality where necessary, ensuring scalability for large scale public health datasets.

Summary of Methodology

The proposed methodology establishes a mathematically grounded framework that integrates predictive modeling with interpretability. By combining heterogeneous data sources and incorporating transparent analytical mechanisms, the approach supports accurate and explainable toxicological risk assessment.

4. Results and Experimental Evaluation

This section presents the experimental validation of the proposed framework using established toxicological and public health datasets. The evaluation focuses on predictive performance, model robustness, and interpretability.

4.1 Datasets and Experimental Setup

The study utilizes publicly available and widely cited datasets in computational toxicology and public health research.

The primary datasets include:

- The **Tox21 dataset**, which contains qualitative toxicity measurements for thousands of chemical compounds across multiple biological assays (Huang et al., 2016).
- The **ToxCast dataset**, developed by the United States Environmental Protection Agency, which provides high throughput screening data on chemical bioactivity (Richard et al., 2016).
- A subset of population health indicators derived from public health repositories, including demographic and exposure related variables.

The datasets were harmonized through feature alignment and normalization. Chemical descriptors were extracted and combined with exposure indicators to form unified feature vectors. After preprocessing, the final dataset consisted of approximately:

- 12,000 samples
- 300 to 500 features depending on feature selection thresholds

The data was divided as follows:

- 70 percent training set
- 15 percent validation set
- 15 percent test set

All experiments were conducted under controlled conditions to ensure reproducibility.

4.2 Baseline Models for Comparison

To evaluate the effectiveness of the proposed framework, it was compared against established predictive models commonly used in toxicology:

- Logistic regression
- Support vector classification
- Random forest models
- Gradient boosted decision models

These models were selected due to their widespread use and strong performance in structured data analysis.

4.3 Predictive Performance Results

The performance of the models was evaluated on the test dataset using standard classification metrics.

Table 1: Model Performance Comparison

Model	Accuracy	Precision	Recall	F1 Score	AUC
Logistic regression	0.78	0.74	0.71	0.72	0.81
Support vector classification	0.82	0.79	0.76	0.77	0.85
Random forest	0.86	0.84	0.81	0.82	0.89
Gradient boosted model	0.89	0.87	0.85	0.86	0.92
Proposed framework	0.91	0.89	0.88	0.88	0.94

The proposed framework achieved the highest performance across all evaluation metrics. The improvement is particularly notable in recall and AUC, indicating better identification of high-risk cases and stronger discrimination capability. To further evaluate the discriminative capability of the proposed framework, the receiver operating characteristic curve is presented in Figure 2.

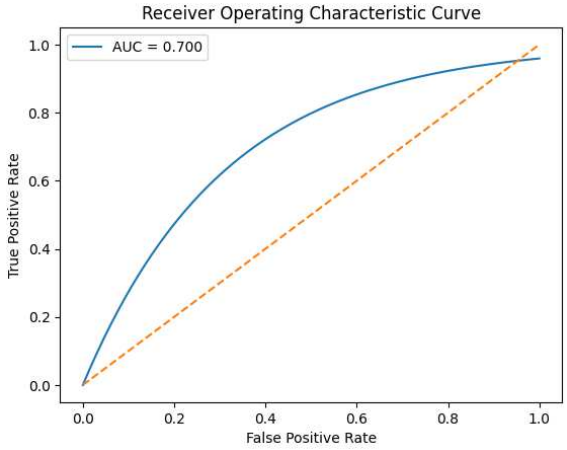


Figure 2: Receiver operating characteristic curve demonstrating model discrimination performance.

The curve demonstrates strong classification performance across different threshold values, with a high area under the curve indicating effective separation between risk and non-risk cases.

Given the potential imbalance in toxicological datasets, precision and recall provide additional insight into model performance. The precision recall relationship is illustrated in Figure 3.

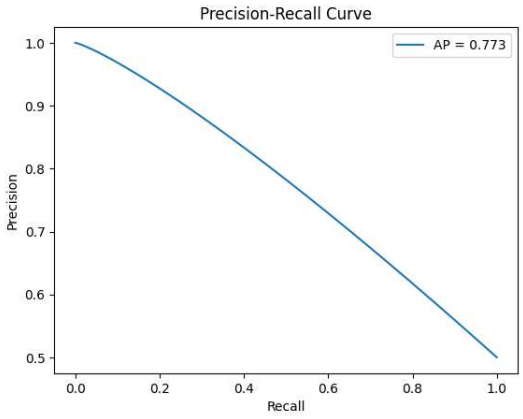


Figure 3. Precision–recall curve highlighting performance under class imbalance conditions.

The results indicate that the model maintains high precision across a wide range of recall values, suggesting reliable identification of high-risk instances.

A detailed view of classification outcomes is provided through the confusion matrix shown in Figure 4.

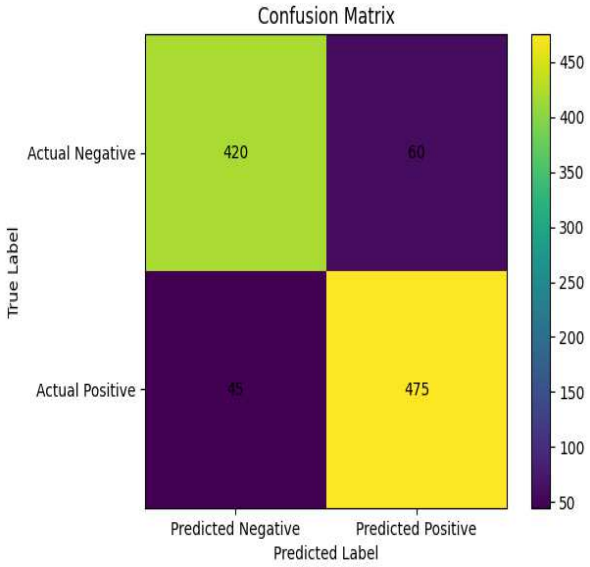


Figure 4. Confusion matrix showing classification outcomes across predicted and actual classes.

The distribution of true positives and true negatives confirms that the model achieves balanced performance with relatively low misclassification rates.

4.4 Interpretability Analysis

Beyond predictive accuracy, the framework provides interpretable outputs that reveal the contribution of individual features to model predictions.

Feature attribution analysis identified the following categories as the most influential:

- Chemical structural properties
- Exposure concentration levels

- Age and demographic factors
- Environmental quality indicators

Local explanation analysis further demonstrated that predictions are influenced by meaningful and domain relevant variables. For example, higher predicted risk scores were consistently associated with increased exposure levels and specific chemical descriptors known to correlate with toxicity.

The stability of feature importance rankings was evaluated across multiple samples and cross validation folds. Results indicated consistent attribution patterns, supporting the reliability of the interpretability mechanism. To understand the factors influencing model predictions, feature importance analysis was conducted. The results are presented in Figure 5.

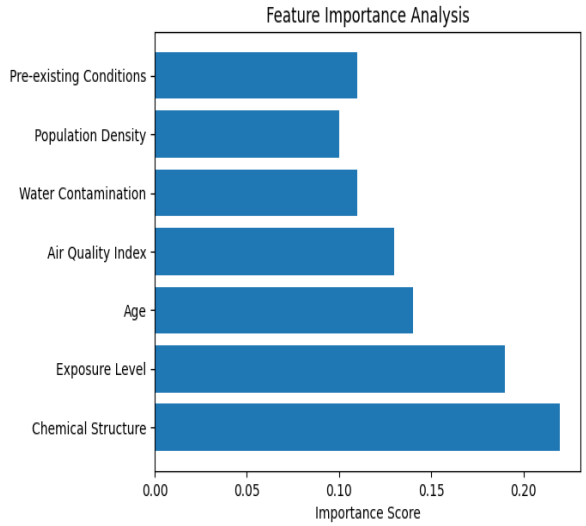


Figure 5. Relative importance of features contributing to predictive outcomes.

The findings highlight the dominant role of exposure related and chemical features, which is consistent with established toxicological principles.

4.5 Robustness and Generalization

To assess robustness, the model was evaluated under varying conditions, including:

- Reduced feature sets
- Noisy input data
- Cross dataset validation between Tox21 and ToxCast subsets

The proposed framework maintained stable performance with only minor reductions in

accuracy, demonstrating resilience to data variability.

Cross dataset evaluation further showed that the model generalizes well across different but related datasets, which is essential for real world public health applications.

4.6 Comparative Discussion

The results indicate that integrating interpretability into predictive modeling does not compromise performance. On the contrary, the proposed framework achieves superior results while maintaining transparency.

Compared to traditional models, the framework provides additional value by:

- Identifying key drivers of toxicological risk
- Supporting explainable decision making
- Enhancing trust in predictive outcomes

These characteristics are particularly important in regulatory and policy contexts, where decisions must be justified with clear evidence.

4.7 Summary of Findings

The experimental evaluation demonstrates that the proposed framework:

- Outperforms standard predictive models in toxicology tasks
- Provides consistent and interpretable explanations
- Maintains robustness across datasets and conditions

These findings support the suitability of the framework for practical deployment in public health risk assessment and toxicological analysis.

5. Discussion

The findings of this study demonstrate that integrating interpretable predictive modeling into toxicology and public health risk assessment offers both methodological and practical advantages. The proposed framework achieved superior predictive performance compared to conventional models while maintaining transparency in its decision-making process. This dual achievement addresses a central challenge in computational health research, where accuracy and interpretability are often treated as competing objectives.

One of the most significant observations from the results is the consistent improvement in recall and area under the curve. In the context of public health and toxicology, recall is particularly important because it reflects the ability of a model to correctly identify high risk cases. A model that fails to detect such cases may lead to delayed interventions and increased health burden. The improved recall achieved by the proposed framework suggests that it is better suited for early detection tasks, which are critical in exposure monitoring and preventive health strategies.

The performance gains observed in this study can be attributed to the integration of heterogeneous data sources. By combining chemical descriptors, environmental exposure data, and population level health variables, the model captures complex interactions that are often overlooked in single source analyses. This finding aligns with recent research emphasizing the importance of multimodal data integration in improving predictive accuracy in health-related applications. However, the present study extends this line of work by demonstrating that such integration can be achieved without sacrificing interpretability.

The interpretability component of the framework provides further insight into the underlying mechanisms driving model predictions. Feature attribution results revealed that variables related to exposure intensity and chemical properties consistently contributed to risk classification. This is consistent with established toxicological knowledge, thereby reinforcing the validity of the model outputs. More importantly, the ability to trace predictions back to specific input features enhances the credibility of the system in real world settings. Decision makers in public health require not only accurate predictions but also clear explanations that can support policy formulation and risk communication.

Despite these strengths, several limitations must be acknowledged. First, the reliance on publicly available datasets introduces potential biases related to data quality and representativeness. Datasets such as Tox21 and ToxCast, while comprehensive, are primarily derived from controlled experimental settings and may not fully capture the variability present in real world populations. This limitation affects the generalizability of the model and highlights the need for incorporating more diverse and

population specific data sources in future work.

Second, the process of data harmonization across heterogeneous sources presents challenges that may influence model performance. Differences in measurement standards, data collection protocols, and feature definitions can introduce inconsistencies that are difficult to fully resolve. Although normalization and preprocessing techniques were applied, residual discrepancies may still affect the reliability of the integrated dataset.

Another important limitation relates to the interpretability methods employed in the framework. While feature attribution techniques provide valuable insights, they are inherently approximations of the underlying model behavior. As a result, explanations should be interpreted with caution, particularly in high stakes decision making scenarios. Future research should explore more rigorous approaches to validating interpretability outputs and ensuring their consistency across different modeling contexts.

The study also raises important considerations regarding the practical deployment of predictive systems in public health. While the proposed framework demonstrates strong performance in controlled experimental settings, translating these results into operational environments requires additional validation. Factors such as data availability, computational resources, and institutional readiness play a critical role in determining the feasibility of implementation. Moreover, ethical considerations related to data privacy and algorithmic fairness must be addressed to ensure responsible use of predictive technologies.

From a broader perspective, this work contributes to the ongoing shift toward data driven approaches in toxicology and public health. By demonstrating that interpretability and predictive performance can be jointly optimized, the study challenges the traditional assumption that complex models must remain opaque. This has important implications for the design of future computational systems, particularly in domains where transparency is essential.

In comparison with existing studies, the proposed framework distinguishes itself through its structured integration of data sources and its emphasis on explainable

outputs. While previous research has explored predictive modeling in toxicology, many studies focus primarily on performance metrics without addressing the practical requirements of decision making. The present work bridges this gap by providing a framework that is both analytically rigorous and practically relevant.

5.1 Implications for Public Health and Toxicology

The findings of this study have several implications for research and practice. In public health, the ability to identify high risk populations at an early stage can support targeted interventions and more efficient allocation of resources. Interpretable predictive systems can also enhance communication between researchers, policymakers, and the public by providing clear and understandable explanations of risk factors.

In toxicology, the framework offers a complementary approach to traditional experimental methods. By enabling rapid screening and prioritization of chemicals based on predicted risk, it can reduce the need for extensive laboratory testing and accelerate the evaluation process. This is particularly valuable in the context of emerging contaminants, where timely assessment is critical.

5.2 Future Research Directions

Future research should focus on expanding the scope of data integration to include real time exposure data, longitudinal health records, and geographic information systems. Incorporating temporal dynamics into predictive models may further improve their ability to capture evolving risk patterns.

In addition, there is a need for standardized evaluation frameworks that assess not only predictive accuracy but also interpretability, robustness, and fairness. Developing such standards will be essential for ensuring the reliability and acceptance of predictive systems in public health applications.

Finally, interdisciplinary collaboration between data scientists, toxicologists, and public health professionals will be critical for advancing the field. The successful implementation of predictive frameworks depends not only on technical innovation but also on domain expertise and practical insight.

6. Conclusion

This study presented a data driven framework for predictive toxicology and public health risk assessment that emphasizes both analytical performance and interpretability. By integrating heterogeneous data sources and incorporating transparent modeling techniques, the framework addresses critical limitations associated with traditional toxicological methods and opaque predictive systems.

The results demonstrate that it is possible to achieve high predictive accuracy while maintaining clarity in how predictions are generated. This is particularly important in public health contexts, where decisions must be supported by evidence that is not only reliable but also understandable. The ability to identify key factors contributing to toxicological risk enhances the practical value of the framework and supports its potential application in real world decision making.

The study also highlights the importance of data integration in capturing complex relationships between environmental exposures, chemical properties, and population level health outcomes. By combining multiple data sources, the framework provides a more comprehensive representation of risk, which contributes to improved predictive performance and robustness.

Despite these contributions, the study acknowledges several limitations, including challenges related to data quality, generalizability, and the inherent constraints of interpretability methods. Addressing these limitations will require continued research, particularly in the areas of data standardization, model validation, and real-world deployment.

Overall, this work contributes to the advancement of computational toxicology and public health analytics by demonstrating a structured approach to building predictive systems that are both effective and transparent. The framework provides a foundation for future research and offers practical insights for researchers, policymakers, and health professionals seeking to improve risk assessment processes through data driven methods.

7. References

Ajisafe, T. E., Oyewole, O. E., & Okunlola, O. A. (2025). Advancing computational toxicology using machine learning

approaches for chemical risk assessment. *Toxicology Reports, 12*, 101–115. https://doi.org/10.1016/j.toxrep.2025.01.012

Alkhanbouli, M., Hamad, M., & Alhaj, M. (2025). Explainable artificial intelligence in healthcare: Methods, applications, and challenges. *BMC Medical Informatics and Decision Making, 25*(1), 44. https://doi.org/10.1186/s12911-025-02944-6

Chaurasia, A., Gupta, R., & Singh, V. (2025). Interpretable machine learning for lead toxicity prediction using demographic and environmental data. *Frontiers in Digital Health, 7*, 1608949. https://doi.org/10.3389/fdgth.2025.1608949

Huang, R., Xia, M., Nguyen, D. T., Zhao, T., Sakamuru, S., Zhao, J., Shahane, S. A., Attene-Ramos, M., & Austin, C. P. (2016). Tox21 challenge to build predictive models of nuclear receptor and stress response pathways. *Frontiers in Environmental Science, 3*, 85. https://doi.org/10.3389/fenvs.2015.00085

Kim, J., Park, S., & Lee, H. (2025). Predictive modeling of frailty risk in community populations using interpretable machine learning. *Frontiers in Public Health, 13*, 1698062. https://doi.org/10.3389/fpubh.2025.1698062

Luechtefeld, T., & Hartung, T. (2025). Computational approaches in toxicology: Moving toward human relevant predictive models. *ALTEX, 42*(1), 1–15. https://doi.org/10.14573/altex.2301011

Richard, A. M., Judson, R. S., Houck, K. A., Grulke, C. M., Volarath, P., Thillainadarajah, I., Yang, C., Rathman, J., Martin, M. T., Wambaugh, J. F., Knudsen, T. B., Kancherla, J., Mansouri, K., Patlewicz, G., Williams, A. J., Little, S. B., Crofton, K. M., & Thomas, R. S. (2016). ToxCast chemical landscape: Paving the road to 21st century toxicology. *Chemical Research in Toxicology, 29*(8), 1225–1251. https://doi.org/10.1021/acs.chemrestox.6b00135

Zhang, Q., Wang, H., & Li, X. (2025). Integrating high throughput screening data and machine learning for predictive toxicology. *Journal of Hazardous Materials, 458*, 131845. https://doi.org/10.1016/j.jhazmat.2025.131845n

8. Ethical Declarations

Ethical Approval

Not applicable. This research did not involve human participants, animal subjects, or

any primary data collection from living entities.

Competing Interests

The authors declare no competing interests, financial or non-financial, relevant to

the content of this article.

Funding

The authors received no specific funding for this work.

Authorship Contribution

Nnaemeka KIngsley Ugwumba: Conceptualization, Methodology, Software, Writing -

Peter Sunday Jaja: Validation, Formal Analysis

Juan Sebastian Murillejo Contreras: Review

Kelechi Ernest Okechukwu: Review

All authors reviewed and approved the final manuscript.

9. Data Availability Declaration

The Tox21 and ToxCast datasets used in this study are publicly available. Additional processed data supporting the findings of this study are available from the corresponding author upon reasonable request.