Multi-Omics Signatures of Oxidative Stress and Inflammation Induced by Short-Term Environmental Particle Gamma Radiation Exposure in Chronic Respiratory Diseases

Sawyer Lehtonen
School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
sawyerwork@uc.edu

Deepeak Riao
Department of Computer Science, Colorado State University, Fort Collins, CO, USA.
raodeepak@colostate.edu

Abstract

The intersection of environmental particle gamma radiation exposure and chronic respiratory disease has emerged as a critical frontier in systems medicine, demanding integrative analytical frameworks capable of resolving complex molecular cascades. Short-term elevation of gamma-emitting radionuclides attached to inhalable particulate matter triggers rapid oxidative stress and systemic inflammation, disproportionately exacerbating conditions such as chronic obstructive pulmonary disease and asthma. This paper presents a large-scale infrastructure perspective on capturing multi-omics signatures—spanning genomics, epigenomics, transcriptomics, proteomics, and metabolomics—that characterize these acute exposure-response dynamics. We articulate a federated data ecosystem architecture that harmonizes high-resolution environmental radiation monitoring with multi-omics profiling deployed across population cohorts. The design confronts fundamental trade-offs between sensor network granularity, biosample temporal resolution, computational scalability, and representation fairness. We examine how such a platform transforms short-term particle gamma activity measurements into actionable signatures through robust AI-driven pattern extraction, while addressing challenges of batch effects, missing modalities, and algorithmic bias that can distort clinical interpretability across diverse subgroups. Governance models grounded in the FAIR principles and federated learning protect participant privacy and data sovereignty, yet raise questions about accountability, consent, and equitable benefit distribution. By situating the omics signatures within a socio-technical infrastructure, the analysis extends to policy instruments that could regulate ambient particle radioactivity, allocate monitoring resources in environmental justice communities, and embed sustainability into longitudinal exposome initiatives. The resulting framework not only illuminates the oxidative and inflammatory pathways linking short-lived radiation exposure to chronic airway damage but also demonstrates how scalable, fair, and resilient system architectures can translate molecular insights into preventive health strategies.

Keywords

multi-omics, oxidative stress, inflammation, particle gamma radiation, chronic respiratory diseases, exposome, machine learning, infrastructure.

1. Introduction

Ambient particulate matter is a complex cocktail of chemical and radioactive constituents, and among the least studied components are the gamma-emitting radionuclides that attach to fine and ultrafine particles. These radionuclides, predominantly progeny of radon-222 and thoron decay, deliver low-dose ionizing radiation to the respiratory epithelium upon inhalation. While chronic radon exposure is a recognized risk factor for lung cancer, the acute pulmonary effects of short-term fluctuations in particle gamma activity—particularly in populations with pre-existing chronic respiratory diseases—remain underexplored at the molecular systems level. A growing body of epidemiological evidence has linked daily changes in ambient particle radioactivity to decrements in lung function, increased rescue medication use, and elevated markers of airway inflammation in individuals with chronic obstructive pulmonary disease (COPD) and asthma [1, 2]. These clinical observations compel a deeper inquiry into the multi-layered biological mechanisms that translate transient radiation exposure into sustained oxidative injury and inflammatory amplification. Multi-omics technologies, which simultaneously assay thousands of molecular features across multiple regulatory layers, offer an unprecedented opportunity to capture the dynamic signatures of such exposure-induced stress responses. However, the realization of this opportunity is not solely a biomedical challenge; it is also a systems engineering and governance challenge, requiring the construction of robust data infrastructures that integrate environmental sensing networks, biospecimen repositories, high-throughput omics pipelines, and interpretable machine learning models in a manner that is scalable, equitable, and sustainable.

The ambition to derive multi-omics signatures of oxidative stress and inflammation from short-term particle gamma radiation exposure in chronic respiratory diseases sits at the confluence of environmental health, computational biology, and large-scale cyberinfrastructure. From a systems perspective, the undertaking demands a departure from siloed, single-layer biomarker studies and toward a holistic, architecture-centric design that treats the entire data-to-knowledge pipeline as an integrated system. This paper analyzes the structural trade-offs inherent in constructing such a system, examining the sensing and data collection modalities, the multi-omics integration frameworks, the AI-driven signature discovery process, and the governance mechanisms required to steer the application of these signatures from benchtop to population health policy. By foregrounding infrastructure, robustness, fairness, and policy implications, we aim to provide an interdisciplinary blueprint that resonates with environmental epidemiologists, systems engineers, and public health decision-makers alike.

2. The Environmental Radiation-Respiratory Health Nexus

The biological initiation linking inhalable particle gamma radiation to airway damage begins with the deposition of radionuclide-carrying particulate matter in the tracheobronchial and alveolar regions. Unlike the well-documented alpha-particle hazard of inhaled radon gas, the gamma radiation emitted by progeny such as bismuth-214 and lead-214 attached to particulate matter delivers penetrating energy that can cause direct ionization of water molecules and biomolecules in epithelial cells and resident immune cells. This energy deposition is spatially non-uniform and temporally episodic, following ambient patterns influenced by meteorology, building ventilation, and outdoor particle dynamics. The resulting short-term bursts of ionizing radiation catalyze the rapid generation of reactive oxygen species (ROS), including superoxide, hydroxyl radicals, and hydrogen peroxide, which overwhelm cellular antioxidant defenses and induce lipid peroxidation, protein carbonylation, and DNA strand breaks [3]. These oxidative lesions are potent activators of the nuclear factor kappa B (NF-κB) and mitogen-activated protein kinase (MAPK) pathways, which orchestrate the transcription of pro-inflammatory cytokines, chemokines, and adhesion molecules. In chronic respiratory diseases such as COPD, where baseline oxidative burden and inflammatory tone are already aberrantly elevated, the superimposed acute insult can precipitate disproportionate epithelial barrier disruption, mucus hypersecretion, and neutrophil-dominated airway influx, manifesting clinically as symptomatic exacerbations.

Population-based studies have begun to characterize the exposure-response relationship for particle gamma radiation and pulmonary outcomes. A panel investigation in COPD patients demonstrated that increases in ambient particle gamma activity over the preceding 24 to 72 hours were associated with significant reductions in forced expiratory volume in one second (FEV1) and forced vital capacity (FVC), independent of particle mass concentration and gaseous co-pollutants [5]. Such findings imply that the radioactive component of particulate matter possesses a toxicological pathway distinct from that of non-radioactive chemical constituents, one that is uniquely capable of driving oxidative modification of respiratory proteins and genomic instability. Complementary work has highlighted that polymorphisms in genes encoding antioxidant enzymes such as glutathione S-transferase and superoxide dismutase modify individual susceptibility to particle-induced lung function decline, suggesting a gene-environment interaction that can be captured through integrated genomic and transcriptomic profiling [4]. Yet the temporal resolution of conventional epidemiological panels, typically sampling outcomes on a weekly or twice-weekly schedule, limits the ability to capture the acute-phase omics responses that peak within hours of exposure and dissipate within a day. To close this gap, monitoring systems must evolve toward high-frequency environmental data collection paired with repeated, just-in-time biospecimen acquisition, creating a data fabric that can resolve the molecular kinetics of oxidative stress and inflammation with unprecedented granularity.

3. Multi-Omics Integration Architecture for Exposure-Response Profiling

Mapping the mechanistic continuum from particle gamma radiation impact to clinical exacerbation requires an integration architecture capable of aligning data streams that differ dramatically in dimensionality, temporal scale, and noise characteristics. Environmental exposure data arrive as continuous time series of gamma activity in various particle size fractions, typically recorded at hourly or sub-hourly intervals by fixed-site radiation monitors or wearable dosimeters. In contrast, multi-omics data are inherently discrete, generated from blood, sputum, nasal lavage, or exhaled breath condensate samples collected at scheduled clinic visits or during acute events. The structural challenge lies in creating a harmonized data model that can associate a rolling window of exposure metrics—mean, peak, and cumulative gamma activity over lags of 0 to 72 hours—with matched profiles of DNA methylation, chromatin accessibility, messenger RNA abundance, protein expression panels, and metabolite fluxes. A modular data lake architecture, underpinned by common data elements and semantic ontologies, provides a scalable solution by decoupling ingestion, transformation, and querying layers. Environmental data are streamed through Apache Kafka pipelines into time-indexed partitions, while omics data are deposited as versioned experimental packages accompanied by extensive metadata conforming to the MIAME and MIAPE standards. A polyglot persistence layer then permits linkage via participant pseudonyms and sample timestamps, enabling temporally resolved association queries without requiring all modalities to reside in a single data store.

The statistical integration itself demands careful navigation of the curse of dimensionality and the heterogeneity of signal types. Penalized regression models, multi-block partial least squares, and Bayesian integrative factor analysis are well-suited to distilling latent variables that explain covariance across the environmental exposure matrix and the stacked omics layers [6]. A landmark framework structures the analysis into sequential stages: first, a block-wise quality control and normalization step that corrects for batch effects and technical drift using housekeeping features and reference samples; second, a sparse canonical correlation analysis that identifies coherent multi-omics modules correlated with the short-term gamma exposure profile; and third, a mediation analysis that tests whether the identified molecular signatures statistically mediate the association between particle gamma radiation and clinical endpoints such as symptom scores or FEV1 decline. By design, this architecture remains agnostic to the specific omics platforms and can accommodate emerging technologies such as single-cell transcriptomics or spatial metabolomics without foundational redesign, as long as the common data model and linking mechanisms are preserved. The resulting multi-omics signatures are not static biomarker panels but dynamic, context-dependent ensembles that capture the interplay between radiation-induced oxidative stress, epigenetic reprogramming of inflammatory genes, and systemic metabolic adjustments.

4. System Design and Data Infrastructure

The deployment of a multi-omics surveillance platform for particle gamma radiation exposure hinges on a resilient and geographically distributed infrastructure that integrates environmental sensing, decentralized biospecimen processing, and secure computational resources. At the environmental layer, a heterogeneous network of regulatory-grade gamma spectrometers, low-cost scintillation detectors, and wearable particulate monitors must be positioned to reflect spatial variability in particle radioactivity across urban, suburban, and industrial gradients. The network's architecture faces a classic resolution-coverage trade-off: high-density monitoring in a few sentinel neighborhoods yields fine-grained exposure surfaces but risks systematic underrepresentation of populations living in environmental justice communities that may face elevated exposures due to proximity to industrial sources or older building stock with higher radon ingress. Mitigating this bias requires a stratified deployment strategy informed by sociodemographic and housing quality data, coupled with adaptive sensor placement algorithms that iteratively optimize coverage for population-weighted exposure misclassification. The sensor data streams are ingested into a cloud-based data integration hub, where quality flags are generated based on detector calibration drift and anomalous radiation readings that could indicate instrument fault rather than true environmental variation.

The biological specimen infrastructure must align with the temporal dynamics of the acute radiation exposure response. A distributed biobanking network that leverages existing COPD and asthma cohort platforms can be augmented with acute sampling protocols: at-home dried blood spot collection, smartphone-prompted saliva sampling during elevated gamma events, and mobile phlebotomy triggered by real-time ambient radiation thresholds. This design transforms the research infrastructure from a passive repository into an active, event-driven system. Specimens are immediately stabilized and transported under cold chain to central omics cores, where high-throughput processing pipelines generate proteomics and metabolomics data within 48 hours of collection, while nucleic acid extracts are preserved for later batch epigenomic and transcriptomic analyses. All data are linked through a privacy-preserving identity management system that employs encrypted pseudonyms and differential privacy mechanisms to prevent re-identification. The federated architecture extends to the computational layer: rather than pooling raw omics data from multiple institutions into a single analytical center—an approach fraught with regulatory hurdles and privacy risks—a federated learning framework allows model parameters to be trained locally and aggregated centrally, preserving data sovereignty while enabling the large sample sizes necessary to detect subtle multi-omics signatures [11]. This system design, though operationally complex, maximizes both scientific validity and ethical acceptability.

5. AI-Driven Signature Discovery: Robustness and Bias Mitigation

Extracting interpretable multi-omics signatures from the integrated data platform necessitates machine learning methodologies that can handle high dimensionality, temporal misalignment, and substantial missingness. Deep learning architectures such as attention-based multimodal transformers and graph neural networks offer the capacity to model complex interactions between time-resolved gamma exposure profiles and the interconnected molecular layers [7, 8]. Attention mechanisms can weight exposure time windows most relevant to downstream gene expression changes, while graph-based representations explicitly encode known protein-protein interaction networks and metabolic pathways as prior knowledge, regularizing the model in ways that enhance both predictive performance and biological plausibility. Despite their power, these models introduce non-trivial robustness concerns. Small perturbations in input exposure data—due to sensor calibration errors or sampling frequency gaps—can propagate through the network and lead to unstable signature compositions. Adversarial training and ensemble approaches that average predictions over multiple model initializations provide practical hardening measures, and model outputs should always be accompanied by uncertainty estimates derived from Monte Carlo dropout or conformal prediction frameworks.

Fairness and bias present equally urgent challenges. Multi-omics reference datasets are overwhelmingly derived from populations of European ancestry, potentially yielding signatures that underperform or mischaracterize oxidative stress pathways in African American, Hispanic, or Indigenous groups who bear disproportionate ambient particle radiation burdens due to systemic housing and occupational inequities. Algorithmic fairness techniques, including reweighting, adversarial debiasing, and subgroup-specific calibration, can be embedded within the training pipeline to reduce performance disparities across self-identified race, ethnicity, and socioeconomic status [9]. However, technical debiasing cannot substitute for intentional recruitment and retention strategies that ensure the underlying data infrastructure reflects the true diversity of exposed populations. Periodic fairness audits, conducted by independent ombuds entities, should inspect the model for differential false positive and false negative rates in predicting exacerbation risk based on the omics signature, and flag populations for whom the signature is insufficiently validated. Furthermore, model interpretability is not merely a methodological nicety but a governance imperative: clinicians and policymakers must understand which molecular pathways are being activated and in which populations, so that targeted prevention strategies can be deployed without inadvertently widening health inequities.

6. Governance, Fairness, and Policy Implications

The translation of multi-omics signatures derived from environmental particle gamma radiation into public health action demands a governance framework that spans data stewardship, regulatory science, and environmental policy. Adopting the FAIR principles—Findable, Accessible, Interoperable, Reusable—for all data assets within the infrastructure is a foundational step, but must be operationalized within a tiered access model that distinguishes between open summary data, registered researcher access to de-identified individual-level omics profiles, and restricted access controlled by community advisory boards for data from sovereign indigenous populations [10, 14]. The federated learning architecture itself serves as a governance instrument, reducing the necessity for central data repositories that could become targets for misuse or breaches. Nonetheless, residual risks of membership inference attacks and model inversion necessitate continuous privacy risk assessment and the adoption of formal privacy guarantees such as differential privacy budgets, which trade a small amount of statistical power for meaningful confidentiality protection [16].

On the regulatory front, particle gamma activity is not currently regulated as a distinct air quality criterion. The evidence base generated by a well-architected multi-omics surveillance system could inform the establishment of short-term ambient gamma radiation guidelines analogous to the World Health Organization's air quality guidelines for particulate matter mass concentration [13]. Establishing such standards would require the integration of omics-derived effect biomarkers into the risk assessment paradigm, shifting the endpoint from mortality or hospital admission counts toward subclinical molecular perturbations that presage future disease burden. This move aligns with the broader exposome concept and with the principles of precision environmental health, where regulation becomes more anticipatory and individualized. Such a shift, however, must be accompanied by equitable resource allocation. Monitoring networks and biobanking facilities are currently concentrated in affluent regions, leaving environmental justice communities—often coincident with zones of higher ambient radioactivity from legacy industrial sites or poorly ventilated housing—without the data necessary to advocate for remediation. Policy instruments such as targeted infrastructure grants, community-based participatory monitoring programs, and radon-resistant building code updates in public housing can leverage multi-omics insights to prioritize interventions precisely where they will yield the greatest reduction in oxidative and inflammatory burden [15, 18].

Sustainability of the infrastructure itself requires a long-term funding model that moves beyond typical five-year research grants. Hybrid models that blend public investment, health system contributions motivated by exacerbation prevention savings, and private sector engagement in sensor innovation and data analytics can provide a durable financial base. The architecture should be designed with open standards and modularity to allow technology evolution without wholesale system replacement. As sensor costs decline and multi-omics assays become more miniaturized and point-of-care, the system could transition from a research platform to a routine environmental health surveillance network, feeding real-time multi-omics risk scores to individuals with chronic respiratory disease via mobile health applications. This progression, while technologically promising, invites careful deliberation about overmedicalization, psychological distress from constant risk awareness, and the potential for discrimination by insurers or employers. Governance must therefore include enforceable anti-discrimination provisions and transparent opt-out mechanisms, ensuring that the system's purpose remains health protection rather than surveillance capitalism.

7. Conclusion

The identification and application of multi-omics signatures of oxidative stress and inflammation triggered by short-term environmental particle gamma radiation represent a paradigmatic opportunity to integrate systems engineering, molecular biology, and public policy. Achieving this integration requires a purpose-built data infrastructure that fuses high-resolution environmental monitoring, event-driven biosample collection, and federated machine learning into a coherent whole. The structural trade-offs between spatial coverage and temporal resolution, between model complexity and interpretability, and between data openness and privacy protection must be navigated deliberately, guided by principles of fairness and sustainability. The derived signatures not only illuminate the pathophysiological cascade from ionizing radiation on particulate matter to airway epithelial injury and inflammatory exacerbation in COPD and asthma but also provide a quantitative substrate for risk assessment and regulatory standard-setting. By embedding these capabilities within a governance framework that centers equity, community participation, and ethical data use, the global health community can harness the digital and biological revolutions to protect those most vulnerable to the often invisible threat of ambient particle radioactivity. The paper has outlined a comprehensive socio-technical blueprint, emphasizing that the ultimate success of multi-omics environmental health research lies as much in the architecture of its infrastructure and the wisdom of its governance as in the molecular signatures it uncovers.

References

1. Brook, R. D., Rajagopalan, S., Pope, C. A., 3rd, Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Jr, Whitsel, L., & Kaufman, J. D. (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation, 121(21), 2331–2378.

2. Darby, S., Hill, D., Auvinen, A., Barros-Dios, J. M., Baysson, H., Bochicchio, F., Deo, H., Falk, R., Forastiere, F., Hakama, M., Heid, I., Kreienbrock, L., Kreuzer, M., Lagarde, F., Mäkeläinen, I., Muirhead, C., Oberaigner, W., Pershagen, G., Ruano-Ravina, A., … Doll, R. (2005). Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case-control studies. BMJ, 330(7485), 223.

3. Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. Genome Biology, 18(1), 83.

4. Spitz, D. R., Azzam, E. I., Li, J. J., & Gius, D. (2004). Metabolic oxidation/reduction reactions and cellular responses to ionizing radiation: a unifying concept in stress biology. Cancer and Metastasis Reviews, 23(3-4), 311–322.

5. Vieira, C. L., Koutrakis, P., Huang, S., Grady, S., Hart, J. E., Coull, B. A., ... & Garshick, E. (2019). Short-term effects of particle gamma radiation activities on pulmonary function in COPD patients. Environmental research, 175, 221-227.

6. Valavanidis, A., Fiotakis, K., & Vlachogianni, T. (2013). Pulmonary oxidative stress, inflammation and cancer: Respirable particulate matter, fibrous dusts and ozone as major causes of lung carcinogenesis through reactive oxygen species mechanisms. International Journal of Environmental Research and Public Health, 10(9), 3886–3907.

7. Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56.

8. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347–1358.

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

10. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018.

11. Liu, Y., Chen, P. C., Krause, J., & Peng, L. (2020). Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Scientific Reports, 10, 12598.

12. Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., Shelow, D., Hindin, D. A., Kilaru, V., & Preuss, P. W. (2013). The changing paradigm of air pollution monitoring. Environmental Science & Technology, 47(20), 11369–11377.

13. Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N., Baldé, A. B., Bertollini, R., Bose-O'Reilly, S., Boufford, J. I., Breysse, P. N., Chiles, T., Mahidol, C., Coll-Seck, A. M., Cropper, M. L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., … Zhong, M. (2018). The Lancet Commission on pollution and health. The Lancet, 391(10119), 462–512.

14. Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. Big Data & Society, 7(2), 2053951720948087.

15. Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructures and their consequences. Sage.

16. Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509–514.

17. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215.

18. Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., & Britter, R. (2015). The rise of low-cost sensing for managing air pollution in cities. Environment International, 75, 199–205.