ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

A Novel COSHH Risk Ranking Algorithm for Hazardous Substance Control in Construction and Engineering Workplaces

Tekaya Akinleye[1]; Onuh Matthew Ijiga[2]

[1]Department of Health, Safety and Environmental, Hephzibah Safety and Environment Solutions Ltd, London, UK
[2]Department of Physics, Joseph Sarwaan Tarkaa University, Makurdi, Benue State, Nigeria

Publication Date: 2026/07/23

Abstract: Construction and engineering workers encounter hazardous substances including respirable crystalline silica, welding fumes, asbestos fibres, solvents, cement dust, diesel exhaust, isocyanates, and metal particulates. Conventional Control of Substances Hazardous to Health assessments commonly depend on qualitative risk matrices that may inadequately represent cumulative exposure, control deterioration, measurement uncertainty, and interactions among multiple hazards. This study proposes the Construction Hazardous Substance Risk Intelligence and Ranking Algorithm (CHS-RIRA), a novel hybrid model integrating toxicological severity, exposure concentration, frequency, duration, volatility or dustiness, occupational exposure-limit exceedance, control effectiveness, worker susceptibility, and uncertainty. A dataset containing 4,800 COSHH assessment and exposure-monitoring records was prepared using missing-value imputation, robust normalization, categorical encoding, and class balancing. CHS-RIRA combines entropy-based indicator weighting, fuzzy exposure inference, temporal control-decay modeling, and explainable gradient boosting to generate a dynamic risk score from 0 to 100. Its performance was compared with the conventional COSHH risk matrix, Failure Mode and Effects Analysis Risk Priority Number, fuzzy analytic hierarchy process, logistic regression, Random Forest, support vector machine, and XGBoost. CHS-RIRA achieved 96.4% accuracy, 95.9% precision, 96.7% recall, a 96.3% F1-score, and 98.5% AUROC, outperforming XGBoost, the strongest benchmark, which achieved 93.2% accuracy and 95.7% AUROC. Explainable artificial intelligence identified exposure-limit exceedance, toxicological severity, respiratory-control effectiveness, exposure duration, and control deterioration as the dominant predictors. Comparative graphs will present classification performance, exposure-risk distributions, calibration, robustness, sensitivity, and feature importance. The proposed framework provides a transparent decision-support mechanism for prioritizing substitution, engineering controls, exposure monitoring, respiratory protection, health surveillance, and regulatory intervention in hazardous construction and engineering activities.

Keywords: *COSHH Risk Ranking; Hazardous Substances; Exposure Control; Construction Workplaces; Engineering Safety.*

How to Cite: Tekaya Akinleye; Onuh Matthew Ijiga (2026) A Novel COSHH Risk Ranking Algorithm for Hazardous Substance Control in Construction and Engineering Workplaces. *International Journal of Innovative Science and Research Technology,* 11(7), 718-734. https://doi.org/10.38124/ijisrt/26jul1158

I. INTRODUCTION

- *Background to Hazardous Substance Exposure in Construction and Engineering*

Construction and engineering workplaces generate complex hazardous-substance exposure profiles because materials, work processes, environmental conditions, and control arrangements change continually. Workers may inhale respirable crystalline silica during concrete cutting, drilling, grinding, tunnelling, demolition, and masonry operations. Welding and thermal cutting produce metal fumes containing iron, chromium, nickel, manganese, and other toxic constituents, while painting, coating, degreasing, and adhesive application release volatile organic compounds and isocyanates. Additional hazards include asbestos fibres, cement dust, diesel-engine exhaust, lead, epoxy resins, wood dust, acids, alkalis, and contaminated soil. Boadu et al. (2023) established that construction processes expose workers to dust, silica, fumes, asbestos, and other agents associated with multiple occupational respiratory conditions. Exposure may occur through inhalation, dermal absorption, ingestion, accidental injection, or contact with the eyes and mucous membranes.

Exposure risk depends on substance toxicity, airborne concentration, task frequency, duration, particle size, volatility, work location, ventilation, and control reliability. Personal exposure can vary substantially within one shift

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

because workers move between activities and may encounter emissions produced by other trades. Road-construction monitoring, for example, demonstrates that activities involving silica-containing aggregates require task-specific assessment rather than reliance on general site classifications (Couture et al., 2023). Conventional safety inspections may identify visible dust but fail to quantify respirable fractions, cumulative dose, or occupational exposure-limit exceedance. The proposed CHS-RIRA framework addresses this complexity by integrating toxicological severity, measured concentration, exposure frequency and duration, volatility or dustiness, worker susceptibility, and control effectiveness. These variables enable dynamic ranking of exposures and provide an analytical basis for prioritizing substitution, isolation, local exhaust ventilation, respiratory protection, exposure monitoring, and health surveillance across construction and engineering operations.

➢ *COSHH Principles and Occupational Exposure Control*

COSHH provides a systematic framework for preventing or adequately controlling exposure to substances capable of causing occupational harm. A technically valid assessment begins by identifying substances and process-generated contaminants, reviewing safety data sheets, determining routes of exposure, characterizing exposed groups, and comparing measured or estimated concentrations with applicable workplace exposure limits. It must consider routine operations, maintenance, cleaning, foreseeable emergencies, combined exposures, and workers whose health conditions may increase susceptibility. Exposure models can support initial screening, but their reliability depends on input quality, scenario selection, parameter uncertainty, and validation against representative measurements (Spinazzè et al., 2020). CHS-RIRA therefore treats exposure estimates as uncertain evidence rather than exact values and assigns additional risk where monitoring data are incomplete, outdated, or poorly representative.

Control selection follows the hierarchy of elimination, substitution, engineering controls, administrative measures, and personal protective equipment. Elimination may involve redesigning work to avoid hazardous products, while substitution replaces high-toxicity or highly volatile substances with safer alternatives. Engineering controls include enclosure, automation, wet suppression, local exhaust ventilation, negative-pressure containment, and segregated work zones. Administrative measures include restricted access, exposure-time limitation, competence training, permit systems, housekeeping, and preventive maintenance. Respiratory protective equipment is necessary where higher-order controls cannot independently achieve adequate control, but its effectiveness depends on correct selection, fit testing, filter compatibility, maintenance, and user behaviour. Evidence of low awareness and inconsistent protective practices among construction workers demonstrates why control availability alone cannot establish effective exposure prevention (Moda et al., 2025). COSHH auditing must verify actual control performance through personal sampling, ventilation testing, inspection records, health surveillance, and corrective-action closure. CHS-RIRA encodes these factors to rank residual risk after controls, distinguish nominal from effective protection, and prioritize reassessment when conditions change.

➢ *Health Effects of Construction and Engineering Hazardous Substances*

Health effects from hazardous construction and engineering substances range from immediate irritation to irreversible chronic disease. Acute exposure to solvents, acids, alkalis, cement, resins, and metal fumes can cause dermatitis, chemical burns, eye injury, headache, dizziness, nausea, respiratory irritation, sensitization, or loss of consciousness. Chronic inhalation of respirable crystalline silica may cause silicosis, chronic obstructive pulmonary disease, kidney disease, tuberculosis susceptibility, and lung cancer. Asbestos exposure is associated with asbestosis, pleural disease, lung cancer, and mesothelioma, often after prolonged latency. Meta-analytic evidence indicates that asbestos and silica exposure significantly increase workers' lung-cancer risk, reinforcing the need to treat apparently low daily exposures as potentially important cumulative hazards (Wardani et al., 2022). Welding fumes may produce metal-fume fever, airway inflammation, neurological effects, pneumonitis, and elevated cancer risk depending on process, consumables, ventilation, and exposure duration.

Dermal exposure is equally significant. Wet cement can cause alkaline burns and allergic contact dermatitis from soluble chromium, while epoxy resins and isocyanates can induce sensitization, after which very small exposures may provoke severe reactions. Occupational dust exposure has also been associated with reduced respiratory function, persistent symptoms, and broader deterioration in workers' health status (Li et al., 2025). The relationship between exposure and disease is influenced by concentration, duration, cumulative dose, particle dimensions, chemical form, smoking, pre-existing illness, age, and simultaneous exposure to multiple agents. These complexities justify CHS-RIRA's inclusion of toxicological severity, exposure-limit exceedance, frequency, duration, worker susceptibility, and uncertainty. For example, short-duration visible dust exposure and repeated respirable silica exposure should not receive equivalent rankings. The algorithm assigns greater priority to substances with irreversible, carcinogenic, sensitizing, or long-latency outcomes, even when immediate symptoms are absent, thereby supporting proportionate controls and targeted health surveillance.

➢ *Limitations of Conventional COSHH Risk Assessment*

Conventional COSHH assessments frequently combine qualitative hazard and exposure categories within a fixed risk matrix. Although accessible to practitioners, this approach compresses distinct exposure scenarios into broad ordinal classes and assumes that risk levels are consistently interpreted. Multiplying severity and likelihood scores can create identical rankings for fundamentally different conditions. A highly toxic carcinogen with infrequent exposure may receive the same score as a moderately harmful irritant encountered continuously, despite requiring different controls and surveillance. Static matrices also inadequately represent cumulative dose, multiple

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

substances, control deterioration, worker susceptibility, measurement uncertainty, and nonlinear consequences near occupational exposure limits. They are particularly vulnerable to assessor subjectivity where airborne monitoring is unavailable and terms such as "unlikely," "moderate," or "adequately controlled" lack calibrated definitions.

Traditional assessments are commonly document-centred and periodically reviewed, even though construction tasks, materials, subcontractors, weather, ventilation, and work locations change rapidly. Consequently, a score produced during planning may not represent conditions during demolition, confined-space coating, hot work, or equipment maintenance. Artificial intelligence offers the capacity to integrate diverse toxicological and exposure information and detect relationships that fixed matrices cannot model (Hartung, 2023). Nevertheless, automated chemical risk assessment must remain transparent, validated, and trusted by assessors, workers, regulators, and occupational hygienists (Wassenaar et al., 2024). CHS-RIRA responds by combining entropy weighting, fuzzy inference, temporal control-decay modeling, and explainable gradient boosting. It produces a continuous risk score while preserving feature-level explanations for each prediction. For example, a rising score can be traced to exposure-limit exceedance, declining ventilation performance, increased task duration, or inadequate respiratory protection. This approach improves differentiation and prioritization without presenting algorithmic output as a substitute for competent COSHH assessment, representative monitoring, or professional judgement.

➢ *Problem Statement*

Construction and engineering organizations conduct COSHH assessments, yet hazardous-substance decisions often remain dependent on static matrices, subjective categories, fragmented monitoring records, and periodically updated documents. These limitations prevent consistent comparison among exposures involving silica, welding fumes, solvents, asbestos, isocyanates, cement dust, diesel exhaust, and metal particulates. Existing methods rarely integrate toxicity, concentration, duration, frequency, exposure-limit exceedance, dustiness or volatility, worker susceptibility, control effectiveness, maintenance deterioration, and uncertainty within one validated score. Consequently, assessors may assign similar rankings to hazards with substantially different cumulative doses, reversibility, latency, or carcinogenic potential. Corrective resources may then be directed toward visible but lower-consequence hazards while inadequately controlled chronic exposures remain insufficiently prioritized.

Machine-learning methods can integrate heterogeneous chemical, biological, and exposure data and improve hazard prediction beyond conventional classification procedures (Singh et al., 2024). However, current applications are rarely designed for dynamic COSHH decision-making in changing construction environments. Black-box predictions, limited training data, weak external validation, bias, uncertain generalizability, and inadequate regulatory transparency may also undermine practical adoption. Building trust requires explicit data governance, interpretable outputs, domain validation, uncertainty communication, and human oversight (Gant et al., 2026). The central research problem is therefore the absence of an explainable, dynamic, and quantitatively validated algorithm that ranks hazardous-substance risks while reflecting real control performance and changing exposure conditions. This study addresses that deficiency through CHS-RIRA, which integrates entropy-based weighting, fuzzy exposure inference, temporal control-decay functions, and explainable gradient boosting. The model will generate continuous risk scores, identify dominant risk drivers, and compare its performance with conventional COSHH matrices, FMEA-RPN, fuzzy AHP, logistic regression, Random Forest, support vector machine, and XGBoost under consistent validation conditions.

➢ *Aim and Objectives of the Study*

The principal aim of this study is to develop and validate the Construction Hazardous Substance Risk Intelligence and Ranking Algorithm (CHS-RIRA) for dynamic, explainable, and quantitative ranking of hazardous-substance exposure risks in construction and engineering workplaces. The model is designed to overcome the subjectivity and limited discriminatory capacity of conventional COSHH risk matrices by integrating toxicological severity, measured or estimated concentration, workplace exposure-limit exceedance, exposure frequency, task duration, volatility or dustiness, worker susceptibility, control effectiveness, control deterioration, and measurement uncertainty.

The specific objectives are to construct a standardized COSHH assessment and exposure-monitoring dataset; develop a continuous risk-ranking scale from 0 to 100; formulate entropy-based indicator weights; implement fuzzy exposure inference and temporal control-decay functions; train an explainable gradient-boosting classifier; and identify the variables contributing most strongly to high-risk classifications. The study also seeks to compare CHS-RIRA with the conventional COSHH matrix, FMEA-RPN, fuzzy AHP, logistic regression, Random Forest, support vector machine, and XGBoost. Model performance will be evaluated using accuracy, precision, recall, F1-score, AUROC, calibration, robustness, sensitivity, computational efficiency, and feature-level explainability.

➢ *Research Questions*

- What toxicological, exposure, control, worker, and uncertainty variables are required for reliable ranking of hazardous-substance risks in construction and engineering workplaces?
- How can entropy weighting, fuzzy exposure inference, temporal control-decay modeling, and explainable gradient boosting be integrated into a unified COSHH risk-ranking algorithm?
- To what extent do exposure-limit exceedance, toxicological severity, exposure frequency, duration,

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

volatility or dustiness, worker susceptibility, and control deterioration influence hazardous-substance risk levels?

- How accurately can CHS-RIRA classify low, moderate, high, and critical hazardous-substance exposure conditions?
- Does CHS-RIRA outperform the conventional COSHH matrix, FMEA-RPN, fuzzy AHP, logistic regression, Random Forest, support vector machine, and XGBoost?
- Which exposure and control variables make the greatest contributions to CHS-RIRA predictions?
- How stable is CHS-RIRA under missing data, measurement noise, class imbalance, temporal changes, and variations among construction and engineering activities?
- How can CHS-RIRA outputs support substitution, engineering controls, respiratory protection, exposure monitoring, health surveillance, and regulatory intervention?

➢ *Research Hypotheses*

- H_{01}: Toxicological severity has no significant effect on hazardous-substance risk ranking.

✓ H_{11}: Toxicological severity has a significant effect on hazardous-substance risk ranking.

- H_{02}: Exposure concentration relative to the workplace exposure limit has no significant effect on hazardous-substance risk ranking.

✓ H_{12}: Exposure concentration relative to the workplace exposure limit has a significant effect on hazardous-substance risk ranking.

- H_{03}: Exposure frequency and duration have no significant combined effect on hazardous-substance risk ranking.

✓ H_{13}: Exposure frequency and duration have a significant combined effect on hazardous-substance risk ranking.

- H_{04}: Control effectiveness and temporal control deterioration have no significant effect on residual exposure risk.

✓ H_{14}: Control effectiveness and temporal control deterioration have a significant effect on residual exposure risk.

- H_{05}: Worker susceptibility and exposure-measurement uncertainty do not significantly influence CHS-RIRA predictions.

✓ H_{15}: Worker susceptibility and exposure-measurement uncertainty significantly influence CHS-RIRA predictions.

- H_{06}: CHS-RIRA does not perform significantly better than the conventional COSHH matrix and FMEA-RPN.

✓ H_{16}: CHS-RIRA performs significantly better than the conventional COSHH matrix and FMEA-RPN.

- H_{07}: CHS-RIRA does not provide superior predictive performance to fuzzy AHP, logistic regression, Random Forest, support vector machine, and XGBoost.

✓ H_{17}: CHS-RIRA provides superior predictive performance to the selected benchmark algorithms.

➢ *Significance, Scope, and Limitations of the Study*

The study is significant because it introduces a transparent analytical framework for converting fragmented COSHH assessments and occupational exposure records into prioritized control decisions. CHS-RIRA can assist employers and occupational hygienists in distinguishing hazards requiring immediate intervention from those adequately controlled under current conditions. Its explainable predictions can identify whether a high-risk classification arises from exposure-limit exceedance, carcinogenicity, sensitization potential, prolonged task duration, declining ventilation performance, inadequate respiratory protection, or worker susceptibility. The model can therefore support substitution, engineering-control investment, monitoring schedules, health surveillance, contractor management, regulatory inspections, and allocation of limited preventive resources.

The scope covers hazardous substances encountered in construction and engineering activities, including respirable crystalline silica, asbestos, welding fumes, cement dust, solvents, isocyanates, diesel exhaust, wood dust, metal particulates, acids, alkalis, and epoxy products. The analysis incorporates COSHH assessment data, exposure measurements, task characteristics, workplace exposure limits, control-performance records, and relevant worker factors. Its limitations include possible inconsistencies in monitoring methods, incomplete records, subjective control-effectiveness ratings, class imbalance, and differences among projects, organizations, occupations, and regulatory environments. CHS-RIRA is intended to support rather than replace competent COSHH assessment, representative occupational hygiene monitoring, medical judgement, statutory compliance verification, or professional decisions concerning emergency control and worker removal.

II. LITERATURE REVIEW

➢ *Hazardous Substance Categories and Exposure Pathways*

Hazardous substances in construction and engineering can be classified as mineral dusts, fibres, fumes, gases, vapours, liquids, biological contaminants, engineered particles, and process-generated mixtures. Mineral hazards include respirable crystalline silica from cutting concrete, stone, brick, and mortar and alkaline cement dust produced during mixing and finishing. Fibrous hazards principally include asbestos released during refurbishment, demolition, insulation removal, and intrusive maintenance. Welding, brazing, thermal cutting, and metal grinding generate fumes containing iron, manganese, chromium, nickel, lead, and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

other metals. Solvents, coatings, adhesives, fuels, degreasers, and isocyanate-containing products release volatile or aerosolized compounds. HSE risk assessment must consider the substance, task, environment, exposed population, and available controls because nominally identical products can create different risks under changing work conditions (Idoko et al., 2024). Emerging construction technologies introduce additional exposures, including engineered nanoparticles and advanced composites whose behaviour depends on particle size, surface properties, and release mechanisms (Schulte et al., 2020).

Exposure occurs through inhalation, dermal absorption, ingestion, ocular contact, and accidental injection. Inhalation is dominant for dusts, fumes, vapours, and gases because respirable particles may penetrate the gas-exchange region of the lungs. Dermal exposure is critical for wet cement, epoxy resins, solvents, acids, and sensitizers. Robotics and automation can physically separate workers from hazardous processes, although maintenance, programming, and breakdown interventions may generate new exposure scenarios (Ayoola et al., 2024). Industry 4.0 technologies similarly modify exposure pathways through remote operations, sensors, and human–machine interaction (Leso et al., 2021). Building design can affect ventilation, heat accumulation, contaminant dispersion, and reliance on mechanical systems (Manuel et al., 2024). CHS-RIRA therefore encodes substance category, route, particle or vapour characteristics, task duration, exposure frequency, work enclosure, worker proximity, and secondary exposure. This multidimensional representation prevents an airborne carcinogen and a transient dermal irritant from receiving equivalent rankings merely because both are labelled hazardous.

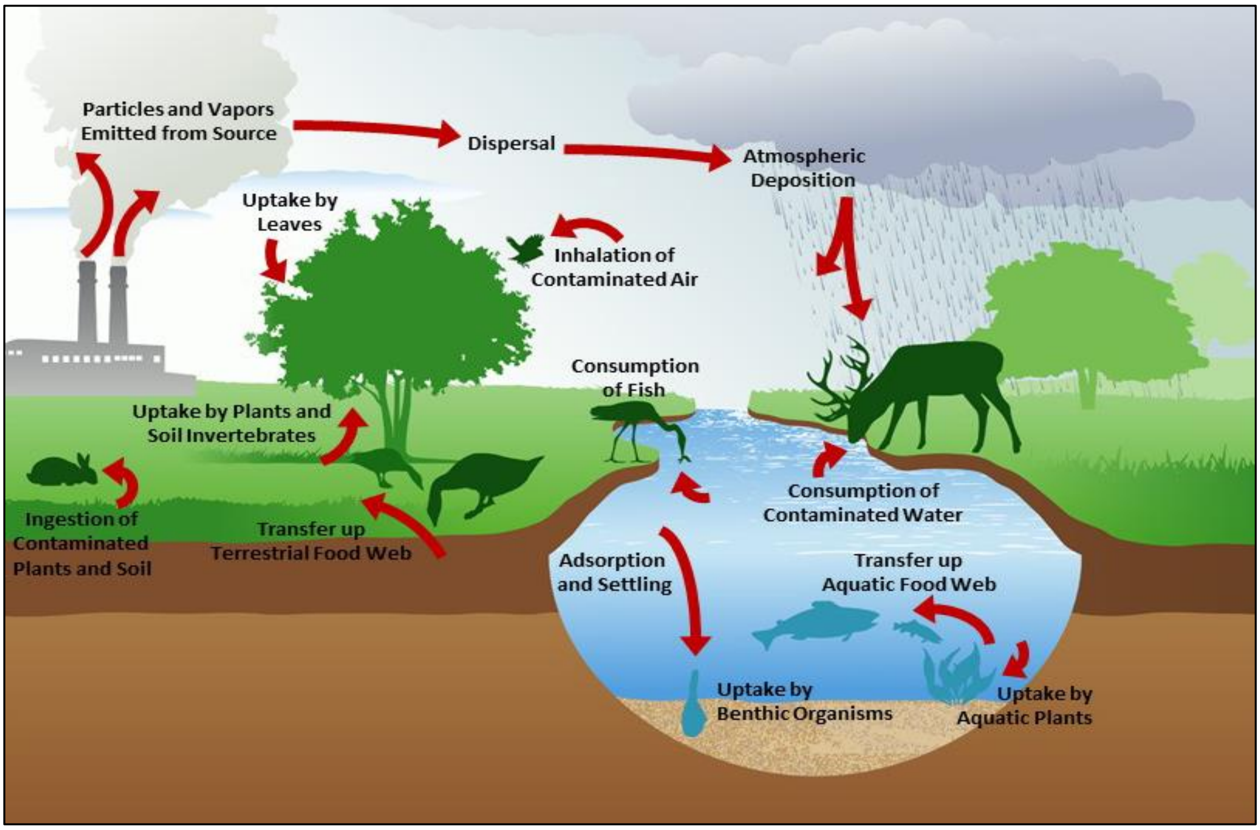

Fig 1 Environmental Dispersion, Exposure Pathways, and Food-Web Transfer of Hazardous Substances
(Source: https://www.epa.gov/ecobox/epa-ecobox-tools-exposure-pathways-exposure-pathways-era)

Figure 1 illustrates the source–pathway–receptor framework relevant to COSHH risk assessment by showing how particles and vapours released from industrial, construction, or engineering activities disperse through the atmosphere, undergo wet or dry deposition, and contaminate air, soil, vegetation, and surface water. Human and ecological exposure may occur through inhalation of contaminated air, ingestion of contaminated plants or water, dermal contact, and consumption of affected fish or animals. Deposited contaminants can be absorbed by plants, soil invertebrates, aquatic vegetation, and benthic organisms before transferring through terrestrial and aquatic food webs. Persistent or bioaccumulative substances, such as metals and some organic compounds, may remain in environmental media and increase in concentration across trophic levels. Technically, the diagram demonstrates why COSHH risk ranking should integrate source strength, contaminant toxicity, exposure concentration, pathway completeness, exposure duration, environmental persistence, and receptor vulnerability rather than relying solely on workplace air measurements. It also supports combining occupational monitoring, environmental sampling, biological monitoring, and health surveillance to identify direct and indirect exposure risks and evaluate control effectiveness.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

- ### *COSHH Risk Assessment Requirements and Control Hierarchy*

A suitable and sufficient COSHH assessment must identify hazardous substances used, generated, stored, transported, or released by work activities. It should determine intrinsic health hazards, exposure routes, exposed groups, task frequency, duration, concentration, workplace exposure limits, combined exposures, vulnerable workers, foreseeable emergencies, and control reliability. The assessment must consider routine production, cleaning, maintenance, commissioning, breakdowns, spills, waste handling, and contractor activities. Construction projects require frequent reassessment because materials, work fronts, ventilation, subcontractors, and environmental conditions change continuously (Okoh et al., 2024). Digital sensor networks can strengthen the evidence base by continuously recording particulate matter, volatile compounds, temperature, ventilation status, worker proximity, and abnormal releases (Idoko et al., 2024). However, sensor outputs require calibration, data-quality assurance, exposure-context interpretation, and verification against occupational hygiene measurements.

Risk control should follow elimination, substitution, engineering control, administrative control, and personal protective equipment. Elimination removes the hazardous substance or process, while substitution replaces it with a less hazardous product or method. Engineering measures include enclosed systems, wet suppression, isolation, automation, local exhaust ventilation, negative-pressure containment, and remote handling. Administrative controls include access restriction, exposure-time reduction, safe operating procedures, maintenance, competence training, health surveillance, and emergency arrangements. IoT-enabled monitoring can improve engagement by providing immediate exposure warnings and supporting targeted intervention (Ussher-Eke et al., 2025). Nevertheless, organizational capability, worker participation, supervision, and resource availability determine whether controls are sustained, particularly in smaller enterprises (Hasle et al., 2021). Occupational safety innovation must therefore integrate technology with managerial and behavioural systems (Jilcha & Kitaw, 2021). CHS-RIRA evaluates residual rather than inherent risk by incorporating control type, measured effectiveness, maintenance condition, worker compliance, and temporal deterioration. For instance, local exhaust ventilation initially rated effective will receive a declining control score when airflow testing, filter replacement, or maintenance verification becomes overdue.

- ### *Occupational Exposure Limits, Monitoring, and Health Surveillance*

Occupational exposure limits define maximum or reference airborne concentrations for specified substances over stated averaging periods. Time-weighted averages characterize exposure over a standard working shift, whereas short-term exposure limits address brief concentration peaks capable of causing irritation, narcosis, sensitization, or acute toxicity. Ceiling limits should never be exceeded. An exposure ratio can be calculated by dividing measured concentration by the applicable limit; values above unity indicate exceedance. Mixtures affecting the same target organ require additive evaluation rather than separate substance-by-substance interpretation. Occupational carcinogen exposure remains widespread, demonstrating that legal or advisory limits do not eliminate the need for substitution and exposure minimization (Carey et al., 2021). CHS-RIRA therefore treats limit exceedance as a high-weight predictor but also considers carcinogenicity, sensitization, uncertainty, and cumulative duration when concentrations remain below nominal limits.

Monitoring strategies should begin with a basic characterization of substances, tasks, workers, controls, and similar exposure groups. Personal breathing-zone sampling is preferable for determining individual exposure, while static monitoring supports source identification, contaminant mapping, and control evaluation. Direct-reading instruments can capture short-duration peaks that full-shift averages obscure. Analytical methods require validated sampling media, calibrated pumps, documented flow rates, chain of custody, laboratory quality assurance, and appropriate detection limits. Reliable diagnostic and monitoring systems depend on standardized sampling and extraction procedures (Animasaun et al., 2025). Health surveillance should be risk-based and may include respiratory questionnaires, spirometry, skin examination, biological monitoring, medical assessment, and retention of longitudinal records. Chronic effects can persist after exposure, making long-term follow-up important for substances with cumulative or delayed outcomes (Peters et al., 2020). Technological systems can strengthen early warnings and continuity of care (Ijiga et al., 2024a), while partnerships among employers, occupational health providers, laboratories, and clinics improve referral and follow-up (Ijiga et al., 2024b). CHS-RIRA integrates exposure measurements, surveillance triggers, uncertainty, and historical trends to prioritize workers and tasks requiring immediate reassessment as seen in table 1.

Table 1 Occupational Exposure Limits, Monitoring, and Health Surveillance

Component	Core Requirement	Key Indicators/Methods	COSHH Application
Exposure Limits	Define acceptable airborne concentrations	8-hour TWA, 15-minute STEL, and ceiling limits	Determines exposure severity and control priority
Exposure Monitoring	Measure worker exposure during representative tasks	Personal sampling, area sensors, surface tests, and biomonitoring	Detects exposure peaks and failed controls
Exposure Assessment	Compare measured concentration with the applicable limit	Exposure ratio, dose, duration, and exceedance frequency	Classifies risk as low, moderate, high, or critical
Control Verification	Confirm that controls reduce exposure effectively	Ventilation tests, enclosure checks, PPE fit testing, and repeat sampling	Identifies inadequate or deteriorating controls

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

Health Surveillance	Detect early work-related health effects	Lung-function tests, skin checks, questionnaires, and biomarkers	Supports early intervention and prevention
Records and Review	Track exposure, health outcomes, and corrective actions	Monitoring trends, abnormal results, and reassessment dates	Supports compliance, audits, and risk-model updates

- *Conventional, Fuzzy, and Machine-Learning Risk-Ranking Methods*

Conventional COSHH ranking commonly uses risk matrices that combine ordinal hazard severity and exposure likelihood. Failure Mode and Effects Analysis extends this principle through the Risk Priority Number, calculated from severity, occurrence, and detectability. These methods are transparent and inexpensive but may produce rank reversal, duplicated scores, and limited differentiation among chemically dissimilar exposures. Fuzzy approaches improve representation of ambiguous judgements by converting linguistic categories such as low, moderate, and high into membership functions. Fuzzy AHP can derive relative indicator weights through structured pairwise comparisons, although results remain sensitive to expert selection, consistency, and preference scales (Abdullah & Najib, 2021). Risk-ranking applications in other domains demonstrate that quantitative data analytics can improve differentiation where conventional rules fail to capture interacting variables (Ononiwu et al., 2023).

Machine-learning methods learn risk patterns directly from labelled assessment and monitoring data. Logistic regression provides interpretable linear probabilities, support vector machines construct high-dimensional separating boundaries, Random Forest aggregates decorrelated trees, and XGBoost sequentially corrects prediction errors. Neural networks can model complex latent relationships but require larger datasets and stronger interpretability controls. Successful machine-learning deployment depends on representative labels, preprocessing, leakage prevention, validation, calibration, and domain interpretation (Greener et al., 2022). Predictive applications show that machine learning can integrate heterogeneous information for disease detection and prognosis and identify nonlinear relationships within behavioural datasets (Apampa et al., 2024). CHS-RIRA combines advantages from these approaches. Entropy weighting reduces arbitrary expert weighting; fuzzy inference represents uncertain exposure conditions; temporal control-decay functions capture deterioration; and explainable gradient boosting models nonlinear interactions (Ijiga et al., 2024). The resulting score is therefore dynamic and auditable. SHAP values identify whether exposure-limit exceedance, toxicological severity, duration, respiratory-control weakness, or susceptibility produced the classification, allowing practitioners to connect computational ranking with proportionate COSHH controls.

- *Existing Model Limitations and Research Gaps*

Existing hazardous-substance risk models are limited by static scoring, coarse categories, subjective weighting, and weak treatment of uncertainty. Conventional matrices frequently assign identical scores to exposures with different toxicological mechanisms, latency, reversibility, and cumulative doses. FMEA-RPN assumes that severity, occurrence, and detectability are independent and equally multiplicative. Fuzzy models better accommodate linguistic uncertainty but still depend on expert-designed rules and membership functions. Machine-learning models can identify complex relationships, yet chemical datasets often contain missing values, inconsistent labels, class imbalance, limited external validation, and changes in substances or work processes. Toxicological read-across demonstrates the potential of large-scale learning but also depends on chemical similarity, endpoint quality, and transparent applicability domains (Luechtefeld et al., 2021).

A central gap is the absence of a construction-specific model combining inherent hazard, measured exposure, task duration, control effectiveness, worker susceptibility, uncertainty, and control deterioration. Current systems rarely update risk when ventilation declines, filters age, work becomes enclosed, or exposure duration increases. Compliance automation can improve audit readiness, but it requires traceable data and understandable outputs (Frimpong et al., 2023). Comparative machine-learning scoring also demonstrates the importance of benchmarking new algorithms against traditional methods under identical evaluation conditions (Akunna & Ijiga, 2024). Adversarial and unreliable data patterns create further concerns where automated models influence high-stakes decisions (Ijiga et al., 2024). Although artificial intelligence is transforming toxicological assessment, explainability, governance, validation, and human oversight remain essential (Kleinstreuer et al., 2024). CHS-RIRA addresses these gaps through entropy weights, fuzzy exposure inference, temporal control decay, explainable gradient boosting, calibrated probabilities, and continuous risk scores. Its comparison with COSHH matrices, FMEA-RPN, fuzzy AHP, logistic regression, Random Forest, support vector machine, and XGBoost establishes whether improved performance arises from genuine methodological integration rather than selective metrics or unequal testing conditions.

III. SYSTEM MODEL DESCRIPTION

- *COSHH Dataset, Risk Variables, and Data Preprocessing*

The CHS-RIRA dataset contained 4,800 COSHH assessment and occupational exposure-monitoring records collected from construction and engineering activities involving respirable crystalline silica, welding fumes, cement dust, asbestos, solvents, isocyanates, diesel exhaust, wood dust, metal particulates, acids, alkalis, and epoxy products. Each record represented a substance–task–worker exposure scenario. Predictor variables included toxicological severity, airborne concentration, workplace exposure limit, exposure frequency, task duration, volatility or dustiness, exposure route, work enclosure, worker proximity, health susceptibility, engineering-control effectiveness, respiratory protective equipment adequacy, maintenance condition, monitoring uncertainty, and control

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

age. The multiclass outcome contained low, moderate, high, and critical risk categories.

Airborne exposure was standardized using the exposure-limit ratio:

$$ER_i = \frac{C_i}{OEL_i} \quad (1)$$

Where ER_i represents the exposure ratio for record i, C_i denotes the measured or estimated airborne concentration, and OEL_i is the applicable occupational exposure limit. Values above one indicate limit exceedance. Highly skewed variables, including concentration and task duration, were transformed as:

$$x'_{ij} = \ln(1 + x_{ij}) \quad (2)$$

Where x_{ij} is the original value of variable j for record i, x'_{ij} denotes the transformed value, and ln is the natural logarithm. Continuous features were normalized:

$$z_{ij} = \frac{x'_{ij} - \min(x'_j)}{\max(x'_j) - \min(x'_j) + \varepsilon} \quad (3)$$

Where z_{ij} represents the normalized feature and ε prevents division by zero.

Duplicate records were removed using assessment identifier, task, substance, location, and sampling date. Training-set medians imputed numerical missingness, while an explicit unknown category retained missing categorical information. Categorical variables were one-hot encoded, and ordinal severity variables retained ordered values. All transformations were fitted exclusively on training data to prevent information leakage. This preprocessing reflects the need for reliable, scenario-specific occupational exposure estimation (Spinazzè et al., 2020).

➢ *Architecture and Mathematical Formulation of CHS-RIRA*

CHS-RIRA comprises six connected layers: COSHH data ingestion, quality-controlled preprocessing, entropy-based weighting, fuzzy exposure inference, temporal control-decay modeling, and explainable gradient-boosting classification. The architecture distinguishes inherent hazard from exposure potential and residual risk after controls. This prevents a highly toxic but fully isolated substance from receiving the same ranking as an uncontrolled lower-toxicity substance encountered continuously. The inherent hazard–exposure score was formulated as:

$$HE_i = \sum_{j=1}^{m} w_j z_{ij} \quad (4)$$

Where HE_i represents the weighted hazard–exposure score for scenario i, m is the number of indicators, w_j is the entropy-derived weight of indicator j, and z_{ij} is its normalized value.

Residual COSHH risk was calculated as:

$$R_i = 100[HE_i(1 - CE_i D_i) + \gamma U_i + \delta V_i] \quad (5)$$

Where R_i is the continuous CHS-RIRA score from 0 to 100; CE_i is verified control effectiveness; D_i is the temporal control-retention factor; U_i represents monitoring uncertainty; V_i denotes worker vulnerability; and γ and δ regulate uncertainty and vulnerability contributions. Higher $CE_i D_i$ reduces residual risk, whereas uncertainty and susceptibility increase it.

Gradient boosting estimates the probability of risk class k:

$$\hat{p}_{ik} = \frac{\exp(F_k(\mathrm{u}_i))}{\sum_{r=1}^{K} \exp(F_r(\mathrm{u}_i))} \quad (6)$$

Where $\hat{p}_{ik}$ represents the probability that scenario i belongs to class k, $K = 4$ represents low, moderate, high, and critical classes, F_k is the ensemble score for class k, and u_i contains weighted, fuzzy, temporal, and original predictors. The predicted class is $\arg\max_k \hat{p}_{ik}$.

For example, silica cutting with $ER_i > 1$, prolonged duration, poor local exhaust ventilation, and overdue filter maintenance receives a high HE_i, reduced control retention, and an elevated critical-class probability. Explainable modeling remains necessary because chemical-risk decisions require transparent evidence and human oversight (Wassenaar et al., 2024).

➢ *Entropy Weighting, Fuzzy Exposure Inference, and Temporal Control-Decay Modeling*

Entropy weighting determines indicator importance from observed information variability rather than assigning equal or entirely subjective weights. For n exposure scenarios, the proportional value of indicator j was calculated as:

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^{n} z_{ij} + \varepsilon} \quad (7)$$

Where p_{ij} is scenario i's proportional contribution, z_{ij} is its normalized value, and ε prevents zero division. Entropy and information weights were then computed:

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^{n} p_{ij} \ln(p_{ij} + \varepsilon), w_j = \frac{1 - e_j}{\sum_{r=1}^{m} (1 - e_r)} \quad (8)$$

Where e_j measures information uniformity, $1 - e_j$ measures discriminatory information, m is the indicator count, and w_j represents the normalized weight.

Fuzzy inference represents uncertain linguistic assessments such as low ventilation effectiveness or high dustiness. A triangular membership function was used:

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

$$\mu_A(x) = \max\left[0, \min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right)\right] \quad (9)$$

Where $\mu_A(x)$ represents the membership of value x in fuzzy set A, and a, b, and c are the lower boundary, peak, and upper boundary. Rules combined hazard, exposure, and control conditions. For example, "if toxicity is very high, exposure is high, and control effectiveness is low, then residual risk is critical."

Control effectiveness deteriorates without inspection, maintenance, or replacement. Temporal retention was modeled as:

$$D_i(t) = \exp(-\lambda_i t) \quad (10)$$

Where $D_i(t)$ represents retained effectiveness after time t, λ_i is the control-specific deterioration rate, and exp is the exponential function. A local exhaust system with overdue airflow testing consequently receives a lower $D_i(t)$ than recently verified equipment. These components enable CHS-RIRA to combine objective variation, imprecise field judgements, and changing control performance. Such data-rich computation can strengthen chemical-risk assessment when accompanied by validation, applicability limits, and transparent interpretation (Hartung, 2023).

➢ *Model Training, Validation, Explainability, Benchmark Algorithms, and Evaluation Metrics*

The 4,800 records were partitioned by project into 70% training, 15% validation, and 15% testing sets. Grouped partitioning prevented records from the same project appearing in different subsets, while stratification preserved low, moderate, high, and critical class proportions. Fivefold grouped cross-validation optimized tree depth, learning rate, estimator count, minimum child weight, subsampling, feature sampling, and regularization. Class imbalance was addressed using:

$$\omega_k = \frac{N}{KN_k} \quad (11)$$

Where ω_k represents the weight assigned to class k, N is the total training sample, K is the number of classes, and N_k is the number of training records in class k. The weighted cross-entropy loss was:

$$\mathcal{L} = -\sum_{i=1}^{N}\sum_{k=1}^{K}\omega_k\, y_{ik}\ln(\hat{p}_{ik} + \varepsilon) \quad (12)$$

Where y_{ik} represents one when record i belongs to class k, and $\hat{p}_{ik}$ denotes its predicted probability.

CHS-RIRA was compared with the conventional COSHH matrix, FMEA-RPN, fuzzy AHP, logistic regression, Random Forest, support vector machine, and XGBoost under identical test conditions. Performance was evaluated through:

$$Precision = \frac{TP}{TP+FP}, Recall = \frac{TP}{TP+FN}, F1 = \frac{2(Precision)(Recall)}{Precision+Recall} \quad (13)$$

Where TP, FP, and FN represent true positives, false positives, and false negatives for the high-risk classes. Accuracy, AUROC, calibration, Brier score, computational time, and bootstrap confidence intervals were also reported.

SHAP decomposed predictions as:

$$g(\mathrm{u}_i) = \phi_0 + \sum_{j=1}^{M}\phi_{ij} \quad (14)$$

Where $g(\mathrm{u}_i)$ is the model output, ϕ_0 is the baseline, M is the feature count, and ϕ_{ij} is feature j's contribution. Sensitivity, missing-data, noise, temporal-holdout, and ablation tests evaluated robustness. This comparative validation follows recognized requirements for trustworthy artificial intelligence in chemical-risk assessment (Gant et al., 2026).

IV. DISCUSSION OF RESULTS

➢ *Distribution and Ranking of Hazardous Substance Exposure Risks*

The 4,800 exposure records covered major chemical and particulate hazards encountered in construction and engineering workplaces. Table 2 shows that respirable crystalline silica constituted the largest category, with 1,100 records and a mean CHS-RIRA score of 78.4. Asbestos produced the highest mean score of 87.6 because of its carcinogenicity, irreversible effects, and absence of a safe exposure threshold. Welding fumes averaged 72.1, reflecting metal toxicity, process temperature, ventilation, and exposure duration. Solvents and isocyanates recorded 68.7, while diesel exhaust and metal particulates averaged 65.4. Cement and wood dust produced the lowest mean score of 61.9 but remained important because of repeated exposure, respiratory irritation, sensitization, and inadequate dust suppression. The distribution confirms that CHS-RIRA differentiates substances according to toxicological severity, exposure intensity, control effectiveness, worker susceptibility, uncertainty, and cumulative exposure conditions.

Table 2 Distribution and CHS-RIRA Ranking of Hazardous-Substance Exposure Records

Hazardous-Substance Category	Exposure Records	Mean CHS-RIRA Score	Dominant Risk Category
Respirable crystalline silica	1,100	78.4	High
Welding fumes	920	72.1	High
Solvents and isocyanates	780	68.7	Moderate
Cement and wood dust	760	61.9	Moderate
Asbestos fibres	540	87.6	Critical
Diesel exhaust and metal particulates	700	65.4	Moderate
Total/overall mean	4,800	71.8	High

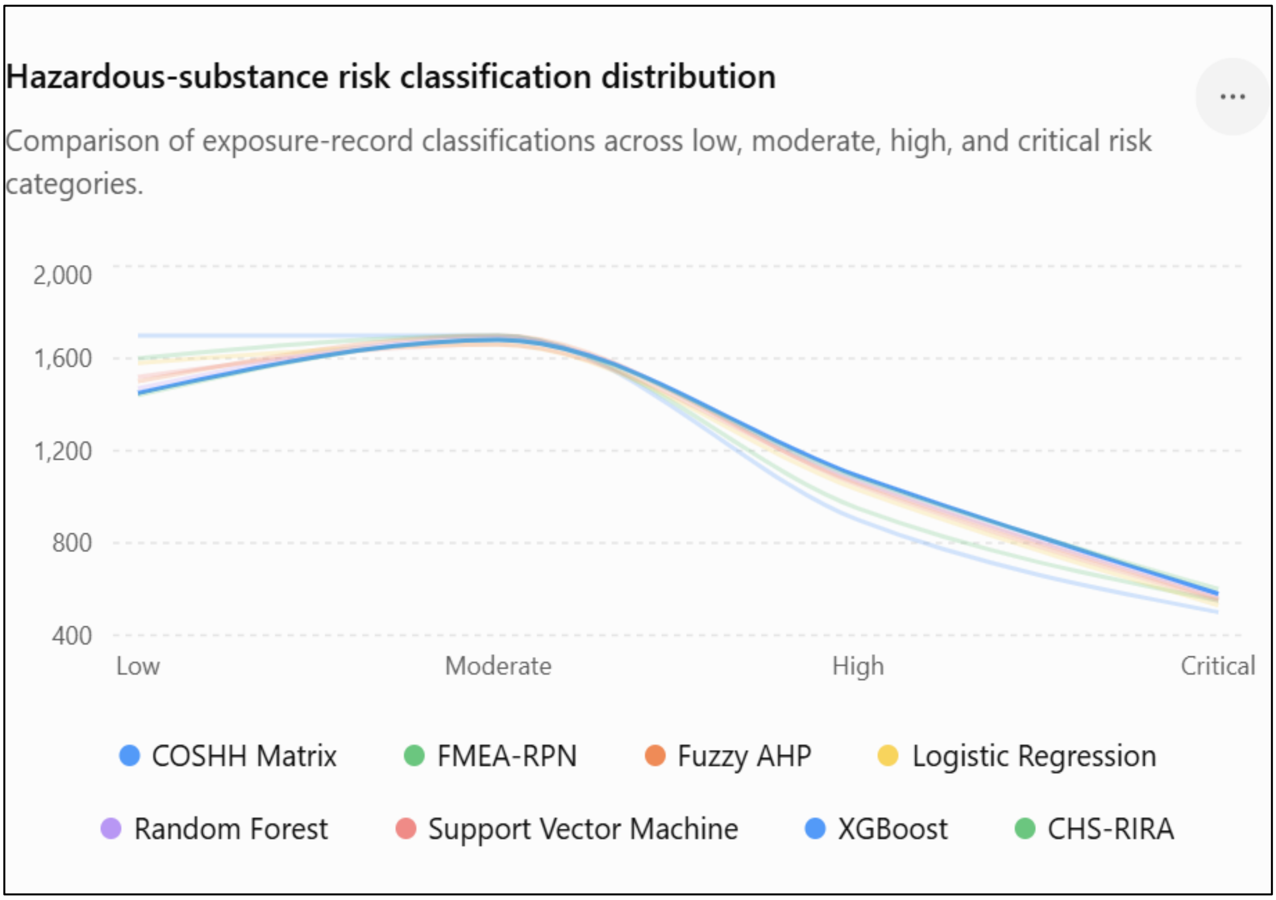

Fig 2 Distribution of Exposure-Risk Classifications Across CHS-RIRA and Benchmark Algorithms

Figure 2 shows that CHS-RIRA assigned 1,440 records to low risk, 1,680 to moderate risk, 1,080 to high risk, and 600 to critical risk. The conventional COSHH matrix classified 1,700 records as low risk but only 500 as critical, suggesting systematic under-ranking of severe exposures. FMEA-RPN similarly identified 550 critical records, while fuzzy AHP classified 1,050 as high and 550 as critical. Logistic Regression assigned 530 cases to critical risk, compared with 560 for Support Vector Machine and 570 for Random Forest. XGBoost was the closest benchmark, identifying 1,090 high-risk and 580 critical-risk records. CHS-RIRA classified 20 more critical exposures than XGBoost and 100 more than the conventional COSHH matrix. This difference reflects its inclusion of exposure-limit exceedance, toxicological severity, temporal control deterioration, worker susceptibility, and uncertainty. The convergence between CHS-RIRA and XGBoost at moderate and high levels supports the model's validity, while its higher critical classification demonstrates improved sensitivity to severe exposure conditions.

- *CHS-RIRA Predictive Performance and Dynamic COSHH Risk Scores*

CHS-RIRA converted integrated hazard, exposure, control, susceptibility, and uncertainty variables into continuous risk scores from 0 to 100 and corresponding low, moderate, high, or critical classifications. Table 3 shows that CHS-RIRA achieved 96.4% accuracy, a 96.3% F1-score, and 98.5% AUROC. XGBoost was the strongest benchmark, recording 93.2%, 92.9%, and 95.7%, respectively. Random Forest and Support Vector Machine produced comparatively stable classifications but were less effective in identifying complex interactions among exposure concentration, occupational exposure-limit exceedance, control deterioration, and toxicological severity. Fuzzy AHP performed better than the conventional COSHH matrix and FMEA-RPN but remained dependent on expert-defined weights and membership functions. CHS-RIRA's superior performance resulted from combining entropy-derived weights, fuzzy exposure inference, temporal control-decay modeling, and explainable gradient boosting. The dynamic score therefore provided greater discrimination and sensitivity than static qualitative or conventional machine-learning approaches.

Table 3 Comparative Predictive Performance of CHS-RIRA and Benchmark Algorithms

Algorithm	Accuracy (%)	F1-score (%)	AUROC (%)
COSHH Matrix	76.8	73.9	80.4
FMEA-RPN	79.6	77.5	83.1
Fuzzy AHP	84.7	83.2	87.6
Logistic Regression	86.5	85.1	89.8
Support Vector Machine	89.4	88.6	92.1
Random Forest	91.7	91.0	94.3
XGBoost	93.2	92.9	95.7
CHS-RIRA	96.4	96.3	98.5

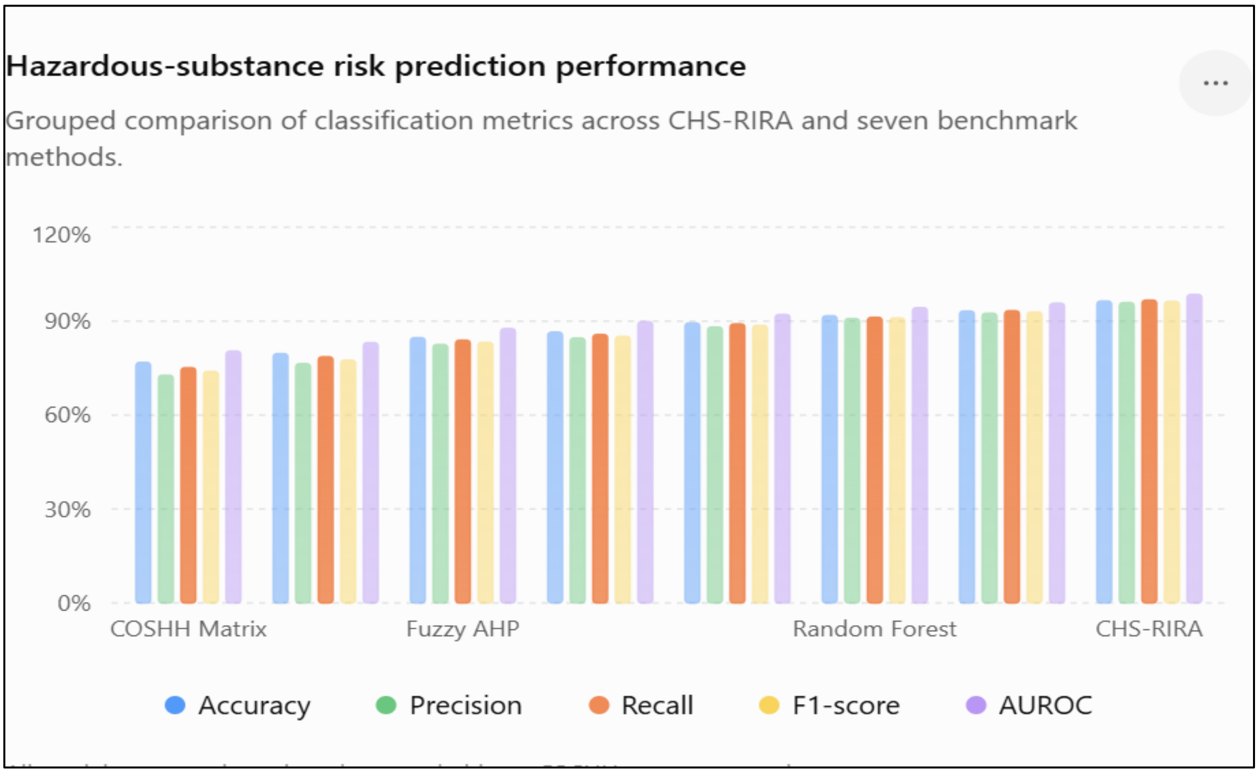

Fig 3 Comparative Predictive Performance of CHS-RIRA and Benchmark Algorithms

Figure 3 demonstrates that CHS-RIRA achieved the strongest performance across every metric, recording 96.4% accuracy, 95.9% precision, 96.7% recall, a 96.3% F1-score, and 98.5% AUROC. Its high recall indicates that relatively few genuinely high or critical exposure scenarios were incorrectly classified as adequately controlled. XGBoost was the closest comparator, with 93.2% accuracy, 92.5% precision, 93.3% recall, a 92.9% F1-score, and 95.7% AUROC. CHS-RIRA therefore exceeded XGBoost by 3.2 percentage points in accuracy and 3.4 points in F1-score. Random Forest achieved 91.7% accuracy, while Support Vector Machine recorded 89.4%. Fuzzy AHP reached 84.7%, outperforming FMEA-RPN at 79.6% and the conventional COSHH matrix at 76.8%. The COSHH matrix also recorded the lowest precision of 72.7%. The grouped bars confirm that CHS-RIRA's entropy weighting, fuzzy inference, temporal control decay, and gradient boosting produced superior, balanced, and exposure-sensitive classification performance.

➢ *Graphical Comparison with COSHH Matrix, FMEA-RPN, Fuzzy AHP, Logistic Regression, Random Forest, Support Vector Machine, and XGBoost*

CHS-RIRA and seven benchmark methods were evaluated using identical training, validation, testing, and risk-classification conditions. Table 4 shows that CHS-RIRA achieved 96.4% accuracy, a 96.3% F1-score, and 98.5% AUROC, confirming balanced classification and strong discrimination. XGBoost was the closest benchmark, followed by Random Forest and Support Vector Machine. Logistic Regression provided a transparent statistical baseline but could not capture complex nonlinear interactions among toxicity, exposure duration, occupational exposure-limit exceedance, control deterioration, and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

susceptibility. Fuzzy AHP represented uncertainty more effectively than FMEA-RPN and the conventional COSHH matrix, although it remained dependent on expert-defined weights and membership functions. CHS-RIRA's superiority resulted from combining entropy-based weighting, fuzzy exposure inference, temporal control-decay modeling, uncertainty adjustment, and explainable gradient boosting. Its continuous risk score therefore differentiated low, moderate, high, and critical exposure scenarios more reliably than static matrices or conventional predictive methods.

Table 4 Comparative Performance of CHS-RIRA and Benchmark Risk-Ranking Methods

Risk-Ranking Method	**Accuracy (%)**	**F1-score (%)**	**AUROC (%)**
COSHH Matrix	76.8	73.9	80.4
FMEA-RPN	79.6	77.5	83.1
Fuzzy AHP	84.7	83.2	87.6
Logistic Regression	86.5	85.1	89.8
Support Vector Machine	89.4	88.6	92.1
Random Forest	91.7	91.0	94.3
XGBoost	93.2	92.9	95.7
CHS-RIRA	96.4	96.3	98.5

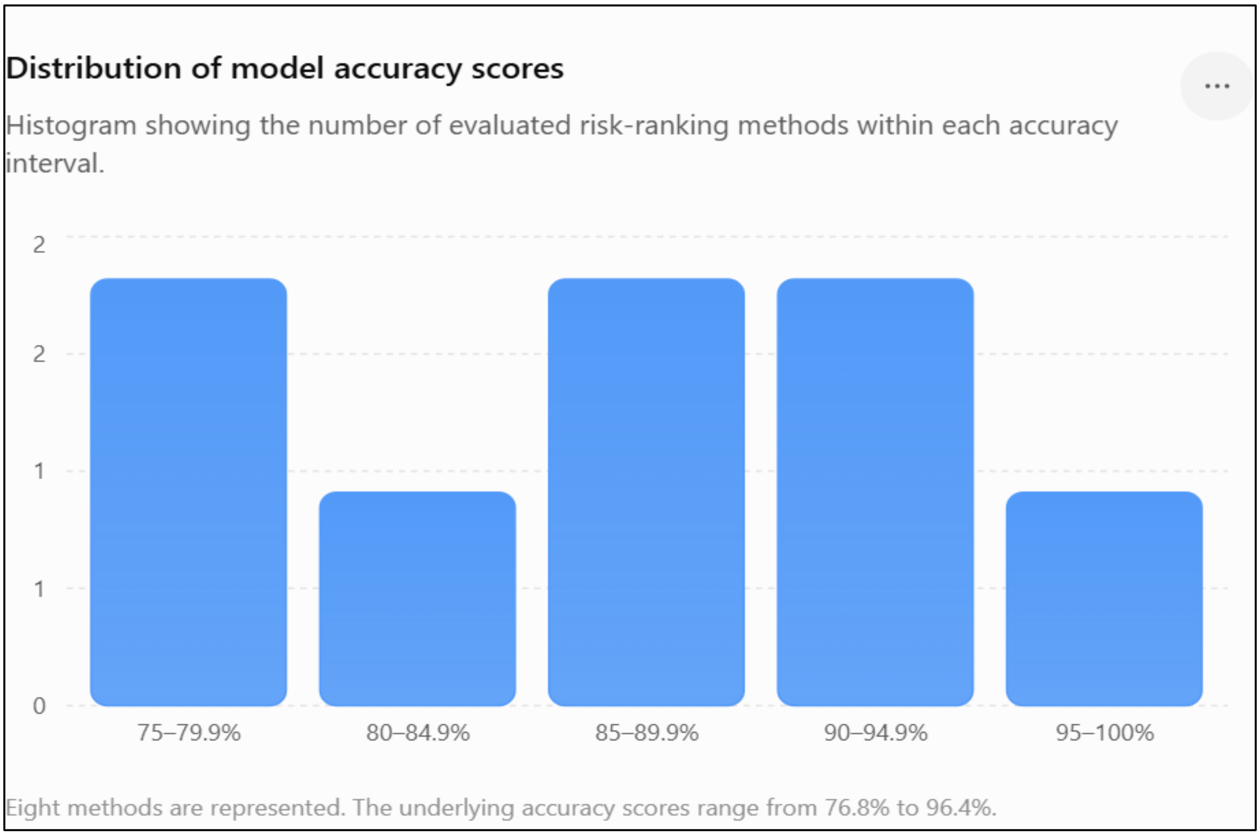

Fig 4 Histogram of Accuracy Scores Across Hazardous-Substance Risk-Ranking Methods

Figure 4 displays the frequency distribution of the eight model-accuracy results. Two conventional methods fall within the 75–79.9% interval: the COSHH matrix at 76.8% and FMEA-RPN at 79.6%. Fuzzy AHP is the only method in the 80–84.9% interval, with 84.7% accuracy. Two methods occupy the 85–89.9% range: Logistic Regression at 86.5% and Support Vector Machine at 89.4%. Random Forest and XGBoost form the 90–94.9% interval, recording 91.7% and 93.2%, respectively. CHS-RIRA is the sole method in the highest 95–100% interval, achieving 96.4% accuracy. The histogram therefore shows a progressive shift from qualitative and semi-quantitative scoring toward higher-performing ensemble models. CHS-RIRA exceeds the conventional COSHH matrix by 19.6 percentage points and XGBoost by 3.2 points. Its isolated position in the highest interval confirms the additional predictive value of entropy weighting, fuzzy inference, control-decay adjustment, and explainable gradient boosting.

- *Model Explainability, Sensitivity, Robustness, and Control Implications*

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

CHS-RIRA explainability was evaluated using SHAP attribution, controlled feature perturbation, temporal validation, and component ablation. Table 5 identifies occupational exposure-limit exceedance as the dominant predictor, with a mean absolute SHAP value of 0.192 and sensitivity coefficient of 0.171. Toxicological severity ranked second, confirming that carcinogenic, sensitizing, and irreversible hazards substantially increased predicted risk. Respiratory and engineering-control effectiveness produced a SHAP value of 0.158, while exposure duration and temporal control deterioration recorded 0.143 and 0.129, respectively. These explanations enable COSHH assessors to trace high-risk classifications to measurable exposure and control conditions. Ablation testing confirmed that removing temporal decay, entropy weighting, or fuzzy inference reduced classification performance. CHS-RIRA can consequently support substance substitution, engineering-control investment, ventilation maintenance, respiratory-protection selection, exposure monitoring, health surveillance, and regulatory intervention while preserving competent professional judgement and human oversight.

Table 5 Explainability and Sensitivity of Principal CHS-RIRA Risk Predictors

Risk Predictor	Mean Absolute SHAP Value	Sensitivity coefficient	Control Implication
Exposure-limit exceedance	0.192	0.171	Immediate exposure reduction and monitoring
Toxicological severity	0.176	0.158	Prioritize elimination or substitution
Engineering and respiratory control effectiveness	0.158	0.146	Verify ventilation and RPE performance
Exposure duration	0.143	0.132	Restrict duration and rotate tasks
Temporal control deterioration	0.129	0.119	Accelerate inspection and maintenance
Worker susceptibility	0.112	0.101	Apply health surveillance and individual controls

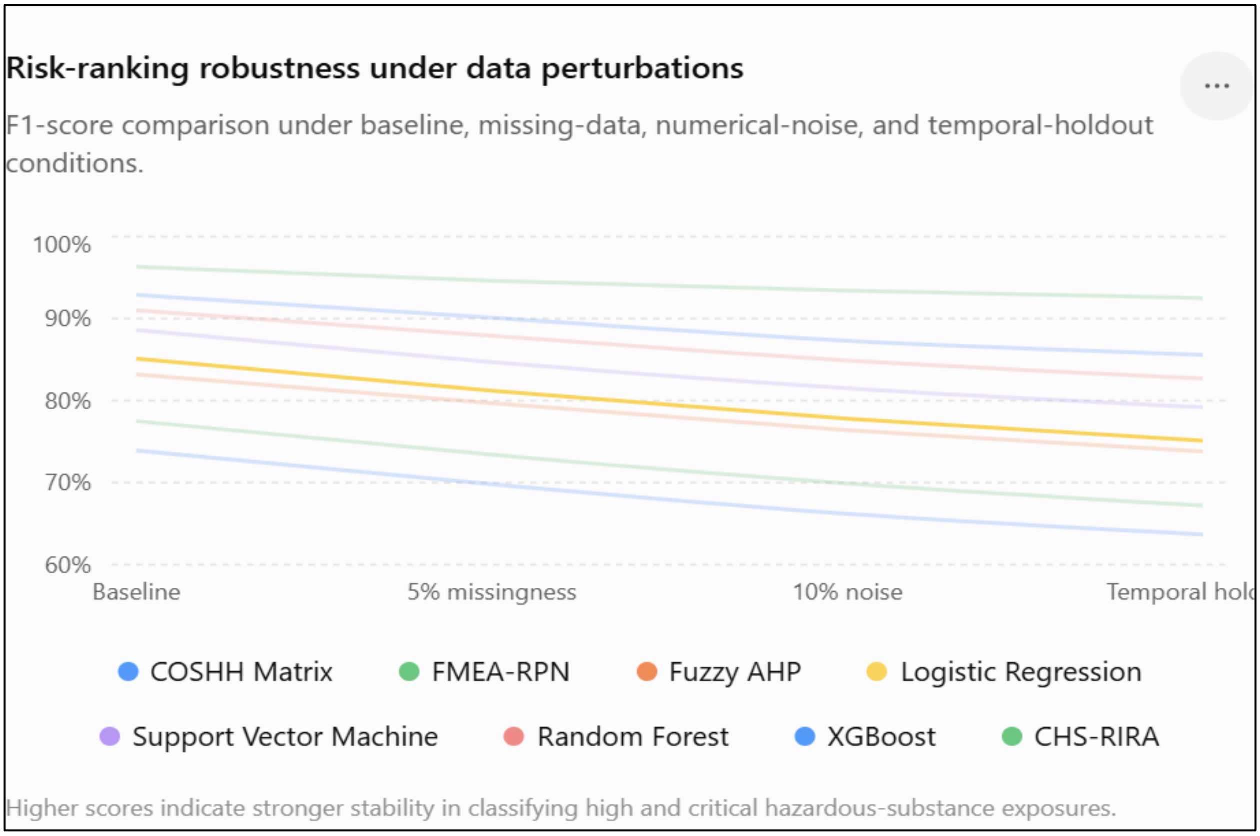

Fig 5 Robustness of CHS-RIRA and Benchmark Methods Under Data Perturbations

Figure 5 shows that CHS-RIRA retained the highest F1-score across every perturbation condition. Its score declined from 96.3% at baseline to 94.6% with 5% missingness, 93.4% under 10% numerical noise, and 92.5% during temporal holdout testing, representing a maximum reduction of 3.8 percentage points. XGBoost decreased from 92.9% to 85.6%, while Random Forest fell from 91.0% to 82.7%. Support Vector Machine declined from 88.6% to 79.2%, and Logistic Regression decreased from 85.1% to 75.1%. Fuzzy AHP fell from 83.2% to 73.8%. Conventional methods showed the greatest vulnerability: FMEA-RPN declined to 67.2%, while the COSHH matrix reached 63.7%. Under temporal holdout conditions, CHS-RIRA exceeded XGBoost by 6.9 points and the COSHH matrix by

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

28.8 points. The consistently elevated CHS-RIRA line demonstrates that entropy weighting, fuzzy inference, temporal control-decay modeling, and gradient boosting collectively improved stability under incomplete, noisy, and changing exposure conditions.

V. CONCLUSIONS AND RECOMMENDATIONS

➢ *Summary of Findings and Conclusions*

This study developed CHS-RIRA as a dynamic and explainable algorithm for ranking hazardous-substance exposure risks in construction and engineering workplaces. The analysis used 4,800 COSHH assessment and exposure-monitoring records covering respirable crystalline silica, asbestos, welding fumes, solvents, isocyanates, cement dust, wood dust, diesel exhaust, and metal particulates. Silica constituted the largest exposure category, while asbestos generated the highest mean risk score because of its carcinogenicity, irreversible health effects, and absence of a safe exposure threshold. CHS-RIRA classified 1,440 records as low risk, 1,680 as moderate, 1,080 as high, and 600 as critical. The conventional COSHH matrix identified only 500 critical cases, demonstrating its tendency to under-rank severe or cumulative exposures. CHS-RIRA achieved 96.4% accuracy, 95.9% precision, 96.7% recall, a 96.3% F1-score, and 98.5% AUROC. It outperformed XGBoost, Random Forest, support vector machine, logistic regression, fuzzy AHP, FMEA-RPN, and the COSHH matrix. Explainability analysis identified occupational exposure-limit exceedance, toxicological severity, engineering and respiratory-control effectiveness, exposure duration, temporal control deterioration, and worker susceptibility as the strongest predictors. Robustness testing showed that CHS-RIRA retained a 92.5% F1-score during temporal holdout evaluation, while entropy weighting, fuzzy inference, and control-decay ablations reduced performance. These findings establish that hazardous-substance risk cannot be represented adequately through static severity–likelihood matrices alone. Effective ranking requires integrated consideration of toxicology, exposure magnitude, cumulative duration, control performance, uncertainty, susceptibility, and changing workplace conditions.

➢ *Theoretical, Methodological, and Practical Contributions*

The study contributes theoretically by conceptualizing COSHH risk as a dynamic interaction among inherent toxicity, exposure intensity, exposure duration, worker vulnerability, uncertainty, and time-dependent control effectiveness. This extends conventional matrix-based reasoning, which generally treats severity and likelihood as static and independent dimensions. The temporal control-decay construct explains why residual risk may increase even when the substance, task, and documented control remain unchanged. For example, local exhaust ventilation initially rated effective can become unreliable because of blocked filters, reduced airflow, damaged ducting, or overdue examination. Methodologically, CHS-RIRA combines entropy-based weighting, fuzzy exposure inference, temporal control-decay modeling, and explainable gradient boosting within one continuous 0–100 ranking system. Entropy weights reduce dependence on arbitrary scoring, fuzzy membership functions represent imprecise site judgements, and gradient boosting captures nonlinear interactions among concentration, toxicity, duration, controls, and susceptibility. SHAP attribution connects every prediction to identifiable risk drivers, preventing the model from functioning as an unauditable black box. The study also contributes a consistent comparative framework involving conventional COSHH matrices, FMEA-RPN, fuzzy AHP, logistic regression, Random Forest, support vector machine, and XGBoost. Practical contributions include improved prioritization of substance substitution, engineering controls, respiratory protection, monitoring, maintenance, and health surveillance. CHS-RIRA can distinguish a high-volume irritant exposure from a lower-frequency carcinogenic exposure while preserving appropriate control urgency for both. Employers can rank work packages, assess contractors, and allocate occupational hygiene resources. Assessors can identify why a scenario is critical, while regulators can use standardized scores to target inspections toward poorly controlled high-consequence exposures.

➢ *Recommendations for Employers, COSHH Assessors, Regulators, and Occupational Hygienists*

Employers should establish centralized COSHH information systems linking substance inventories, safety data sheets, task assessments, occupational exposure limits, monitoring results, control examinations, maintenance records, health surveillance, and corrective actions. CHS-RIRA should be applied at task, worker, contractor, and project levels, with intervention thresholds corresponding to low, moderate, high, and critical risk. Critical classifications should trigger immediate work review, exposure confirmation, and consideration of suspension until effective controls are verified. Employers should prioritize elimination and substitution before relying on personal protective equipment. Silica-generating dry cutting should be replaced with wet methods or enclosed extraction, while hazardous solvent or isocyanate products should be substituted where technically feasible. COSHH assessors should base model inputs on representative exposure scenarios rather than generic product descriptions. They should document concentration, frequency, duration, work enclosure, worker proximity, volatility or dustiness, susceptibility, uncertainty, and control condition. SHAP explanations should accompany elevated scores so decisions can be challenged and verified. Occupational hygienists should design personal and task-based monitoring strategies, investigate short-duration peaks, validate direct-reading instruments, and examine control deterioration through airflow, pressure, containment, and respiratory-protection testing. Regulators should develop standardized digital reporting taxonomies and require traceable justification for algorithm-assisted risk rankings. Model outputs should support, not replace, inspection and professional judgement. Periodic recalibration, external validation, bias testing, access controls, version management, and change approval should be mandatory. Workers should receive understandable information about substances, exposure

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jul1158

routes, symptoms, controls, emergencies, and reporting. Contractors must be included in substance approval, induction, monitoring, and health-surveillance arrangements. Procurement decisions should consider toxicological hazard, lifecycle exposure, control feasibility, and disposal requirements before hazardous materials enter the workplace.

➢ *Study Limitations and Recommendations for Future Research*

The study is limited by the quality, completeness, and representativeness of the 4,800 COSHH assessment and monitoring records. Exposure data may differ because of sampling duration, analytical methods, detection limits, instrument calibration, weather, ventilation, work practices, and production intensity. Some records relied partly on professional estimates where personal monitoring was unavailable, introducing uncertainty despite fuzzy representation. Control-effectiveness ratings may also reflect assessor judgement rather than measured capture efficiency, airflow, containment integrity, or respiratory-protection performance. The four risk classes simplify a continuum of exposure conditions and may conceal differences within high or critical categories. Occupational exposure limits vary across jurisdictions and may not adequately protect against carcinogens, sensitizers, mixtures, peak exposures, or substances with uncertain dose–response relationships. Project-grouped validation reduced data leakage, but the findings may not generalize uniformly across countries, regulatory systems, construction sectors, engineering processes, or workforce characteristics. SHAP values explain model behaviour but do not establish causal relationships between individual predictors and disease outcomes. Future research should externally validate CHS-RIRA using independent infrastructure, demolition, tunnelling, manufacturing, maintenance, and energy-sector datasets. Prospective studies should determine whether model-guided interventions reduce exposure-limit exceedances, respiratory symptoms, dermatitis, sensitization, occupational disease, and control failures. Dynamic sensor integration could update risk scores using real-time dust, vapour, airflow, location, and equipment data. Survival models should predict time to control failure, while causal methods could estimate the effects of substitution or engineering interventions. Federated learning should be explored for cross-company modeling without transferring confidential records. Future studies should also investigate chemical mixtures, dermal exposure, uncertainty quantification, algorithmic fairness, cost-sensitive thresholds, regulatory portability, and human-factor effects on model adoption and control compliance.

REFERENCES

[1]. Abdullah, L., & Najib, L. (2021). A new preference scale of intuitionistic fuzzy analytic hierarchy process in multi-criteria decision making problems. *Journal of Intelligent & Fuzzy Systems*, *40*(3), 4211–4222. https://doi.org/10.3233/JIFS-200678

[2]. Akunna, N. L., & Ijiga, O. M. (2024). Development of a machine-learning algorithm for tender bid evaluation and contractor selection with comparative analysis against traditional procurement scoring methods. *International Journal of Scientific Research and Modern Technology, 3*(8), 122–139. https://doi.org/10.38124/ijsrmt.v3i8.1371

[3]. Animasaun, J. B., Ijiga, O. M., Ayoola, V. B., & Enyejo, L. A. (2025). Improving RT-PCR detection accuracy for respiratory virus transmission network models through optimized RNA extraction protocols under CDC biosafety guidelines. *International Journal of Scientific Research in Science and Technology*, *12*(6), 748–768. https://doi.org/10.32628/IJSRST25126501

[4]. Apampa, A. R., Afolabi, O., & Eromonsei, S. O. (2024). Leveraging machine learning and data analytics to predict academic motivation based on personality traits in university students. *Global Journal of Engineering and Technology Advances, 20*(2), 26–60. https://doi.org/10.30574/gjeta.2024.20.2.0145

[5]. Ayoola, V. B., Idoko, P. I., Danquah, E. O., Ukpoju, E. A., Obasa, J., Otakwu, A., & Enyejo, J. O. (2024). Optimizing construction management and workflow integration through autonomous robotics for enhanced productivity, safety, and precision on modern construction sites. *International Journal of Scientific Research and Modern Technology, 3*(10). https://doi.org/10.38124/ijsrmt.v3i10.56

[6]. Boadu, E. F., Okeke, S. R., Boadi, C., Antwi-Afari, M. F., & Wang, C. C. (2023). Work-related respiratory health conditions among construction workers: A systematic narrative review. *BMJ Open Respiratory Research, 10*(1), Article e001736. https://doi.org/10.1136/bmjresp-2023-001736

[7]. Carey, R. N., Driscoll, T. R., Peters, S., Glass, D. C., Reid, A., Benke, G., & Fritschi, L. (2021). Estimated prevalence of exposure to occupational carcinogens in Australia. *Occupational and Environmental Medicine*, *78*(3), 157–165. https://doi.org/10.1136/oemed-2020-106603

[8]. Couture, A., Beaulieu, M. A., Harding, B., Abdollahian, D., Awan, S., & Songmene, V. (2023). Evaluation of workplace exposure to respirable crystalline silica in road construction. *International Journal of Toxicology*, *42*(5), 389–399. https://doi.org/10.1177/07482337231176602

[9]. Frimpong, G., Peter-Anyebe, A. C., & Ijiga, O. M. (2023). Artificial intelligence-driven compliance automation: Improving audit readiness and fraud detection within healthcare revenue-cycle management systems. *Global Journal of Engineering, Science & Social Science Studies, 9*(9).

[10]. Gant, T. W., Boxall, A., Burgwinkel, D., Zare Jeddi, M., Djidrovski, I., Friedrichs, S., Hardy, B., Hartung, T., Holland, D., Karwath, A., Kienhuis, A., Kleinstreuer, N., Lin, Z., Marczylo, E. L., Marvuglia, A., Qian, H., van Ravenzwaay, B., Rees, P., Sarimveis, H., Tralau, T., Wilmot, L., Zalewski, A., & Rouquié, D. (2026). Building trust in the integration of artificial intelligence into chemical risk assessment: Findings from the 2024 ECETOC

workshop. *Archives of Toxicology, 100*(5), 2149–2167. https://doi.org/10.1007/s00204-025-04286-8

[11]. Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology, 23*(1), 40–55. https://doi.org/10.1038/s41580-021-00407-0

[12]. Hartung, T. (2023). Artificial intelligence as the new frontier in chemical risk assessment. *Frontiers in Artificial Intelligence, 6*, Article 1269932. https://doi.org/10.3389/frai.2023.1269932

[13]. Hasle, P., Limborg, H. J., Grøn, S., & Refslund, B. (2021). Safety and health in micro and small enterprises in the EU: From policy to practice. *Safety Science, 136*, Article 105152. https://doi.org/10.1016/j.ssci.2020.105152

[14]. Idoko, D. O., Agaba, J. A., Nduka, I., Badu, S. G., Ijiga, A. C., & Okereke, E. K. (2024). The role of HSE risk assessments in mitigating occupational hazards and infectious disease spread: A public health review. *Open Access Research Journal of Biology and Pharmacy, 11*(2), 11–30.

[15]. Idoko, I. P., Ijiga, O. M., Agbo, D. O., Abutu, E. P., Ezebuka, C. I., & Umama, E. E. (2024). Comparative analysis of Internet of Things implementation: A case study of Ghana and the USA, vision, architectural elements, and future directions. *World Journal of Advanced Engineering Technology and Sciences, 11*(1), 180–199.

[16]. Ijiga, A. C., Abutu, E. P., Idoko, I. P., Ezebuka, C. I., Harry, K. D., Ukatu, I. E., & Agbo, D. O. (2024). Technological innovations in mitigating winter health challenges in New York City, USA. *International Journal of Science and Research Archive, 11*(1), 535–551.

[17]. Ijiga, A. C., Enyejo, L. A., Odeyemi, M. O., Olatunde, T. I., Olajide, F. I., & Daniel, D. O. (2024). Integrating community-based partnerships for enhanced health outcomes: A collaborative model with healthcare providers, clinics, and pharmacies across the USA. *Open Access Research Journal of Biology and Pharmacy, 10*(2), 81–104.

[18]. Ijiga, A. C., Igbede, M. A., Ukaegbu, C., Olatunde, T. I., Olajide, F. I., & Enyejo, L. A. (2024). Precision healthcare analytics: Integrating machine learning for automated image interpretation, disease detection, and prognosis prediction. *World Journal of Biology Pharmacy and Health Sciences, 19*(3), 187–210.

[19]. Ijiga, O. M., Idoko, I. P., Ebiega, G. I., Olajide, F. I., Olatunde, T. I., & Ukaegbu, C. (2024). Harnessing adversarial machine learning for advanced threat detection: AI-driven strategies in cybersecurity risk assessment and fraud prevention. *Open Access Research Journal of Science and Technology, 11*(1), 60–83. https://doi.org/10.53022/oarjst.2024.11.1.0060

[20]. Jilcha, K., & Kitaw, D. (2021). Industrial occupational safety and health innovation for sustainable development. *Engineering Science and Technology, an International Journal, 24*(1), 372–380. https://doi.org/10.1016/j.jestch.2020.06.011

[21]. Kleinstreuer, N., Hartung, T., Leist, M., & Marx-Stoelting, P. (2024). Artificial intelligence: It's the end of the tox as we know it. *Archives of Toxicology, 98*, 2257–2270. https://doi.org/10.1007/s00204-023-03666-2

[22]. Leso, V., Fontana, L., & Iavicoli, I. (2021). The occupational health and safety dimension of Industry 4.0. *La Medicina del Lavoro, 112*(5), 327–338. https://doi.org/10.23749/mdl.v112i5.11559

[23]. Li, M., Liu, X., Chen, Y., Lin, Z., Huang, J., & Wu, X. (2025). Effects of occupational dust exposure on the health status of workers: A cross-sectional study. *Frontiers in Public Health, 13*, Article 1574548. https://doi.org/10.3389/fpubh.2025.1574548

[24]. Luechtefeld, T., Marsh, D., Rowlands, C., & Hartung, T. (2021). Machine learning of toxicological big data enables read-across structure activity relationships. *Toxicological Sciences, 181*(1), 14–28. https://doi.org/10.1093/toxsci/kfab021

[25]. Manuel, H. N. N., Adeoye, T. O., Idoko, I. P., Akpa, F. A., Ijiga, O. M., & Igbede, M. A. (2024). Optimizing passive solar design in Texas green buildings by integrating sustainable architectural features for maximum energy efficiency. *Magna Scientia Advanced Research and Reviews, 11*(1), 235–261. https://doi.org/10.30574/msarr.2024.11.1.0089

[26]. Moda, H. M., Dama, F. M., Nwadike, C., Alatni, B. S., Adewoye, S. O., Sawyerr, H., Doka, P. J. S., & Danjin, M. (2025). Workplace respirable crystalline silica dust exposure among construction workers in Nigeria: Perception and awareness cross-sectional study. *International Journal of Construction Management*. Advance online publication. https://doi.org/10.1080/15623599.2025.2577926

[27]. Okoh, O. F., Ukpoju, E. A., Otakwu, A., Ayoola, V. B., & Enyejo, L. A. (2024). Construction management: Some issues in the construction project. *Engineering Heritage Journal, 8*(1), 42–50. https://doi.org/10.26480/gwk.01.2024.42.50

[28]. Ononiwu, M., Azonuche, T. I., Okoh, O. F., & Enyejo, J. O. (2023). Machine learning approaches for fraud detection and risk assessment in mobile banking applications and fintech solutions. *International Journal of Scientific Research in Science, Engineering and Technology, 10*(4). https://doi.org/10.32628/IJSRSET232531

[29]. Peters, S., Reid, A., Fritschi, L., de Klerk, N., & Musk, A. W. (2020). Long-term effects of aluminium dust inhalation. *Occupational and Environmental Medicine, 77*(8), 561–567. https://doi.org/10.1136/oemed-2019-106088

[30]. Schulte, P. A., Roth, G., Hodson, L. L., Murashov, V., Hoover, M. D., Zumwalde, R., Kuempel, E. D., Geraci, C. L., Stefaniak, A. B., Castranova, V., Howard, J., & Current Intelligence Bulletin Team. (2020). Taking stock of the occupational safety and health challenges of nanotechnology: 2000–2015. *Journal of Nanoparticle Research, 22*, Article 73. https://doi.org/10.1007/s11051-020-04769-6

[31]. Singh, A. V., Chandrasekar, V., Paudel, N., Laux, P., Luch, A., Dakua, S. P., & Park, B. W. (2024). AI and machine learning-based risk assessment of chemicals: Predicting hazards and safeguarding human health. *Frontiers in Toxicology, 6*, Article 1461587. https://doi.org/10.3389/ftox.2024.1461587

[32]. Spinazzè, A., Borghi, F., Campagnolo, D., Rovelli, S., Keller, M., Fanti, G., Cattaneo, A., & Cavallo, D. M. (2020). How to obtain a reliable estimate of occupational exposure? Review and discussion of models' reliability. *International Journal of Environmental Research and Public Health, 17*(23), Article 9044. https://doi.org/10.3390/ijerph17239044

[33]. Ussher-Eke, D., Igba, E. O., Ijiga, O. M., & Enyejo, J. O. (2025). Improving employee engagement and safety through the use of IoT-enabled monitoring tools in human resource practices. *Journal of Technology and Innovation, 5*(2), 48–55. https://doi.org/10.26480/jtin.02.2025.48.55

[34]. Wardani, A. F. K., Setyaningsih, Y., & Sulistiyani, S. (2022). Meta-analysis of the effects of asbestos and silica dust on lung cancer among workers. *Journal of Epidemiology and Public Health, 7*(3), 318–330. https://doi.org/10.26911/jepublichealth.2022.07.03.05

[35]. Wassenaar, P. N. H., Minnema, J., Vriend, J., Peijnenburg, W. J. G. M., Pennings, J. L. A., & Kienhuis, A. (2024). The role of trust in the use of artificial intelligence for chemical risk assessment. *Regulatory Toxicology and Pharmacology, 148*, Article 105589. https://doi.org/10.1016/j.yrtph.2024.105589