scientific reports

Check for updates

OPEN

# Applying machine learning algorithms to explore the impact of combined noise and dust on hearing loss in occupationally exposed populations

Yong Li[1,2]✉, Xin Sun[3], Yongtao Qu[2], Shuling Yang[4], Yueyi Zhai[4] & Yan Qu[1,5]✉

**This study aimed to explore the combined impacts of occupational noise and dust on hearing and extra-auditory functions and identify associated risk factors via machine learning techniques. Data from 14,145 workers (627 with occupational noise-induced hearing loss (ONIHL)) at Hebei Medical Examination Center (2017–2023) were analyzed. Workers with combined exposure and without specific contraindications or other hearing impairment causes were included. Demographic and clinical data were gathered. Chi-square and Mann-Whitney U tests examined variables, and multivariate logistic regression determined ONIHL risk factors. Machine learning algorithms like Logistic Regression and Random Forest were developed, optimized, and evaluated. Results showed significant differences in gender, exposure, blood pressure, smoking, etc. between ONIHL and non-ONIHL groups. Male gender, combined exposure, diastolic blood pressure elevation, smoking, fasting blood glucose elevation, and age were positive predictors, while systolic blood pressure elevation was negative. The logistic model had the highest predictive ability (ROC = 0.714). Subgroup analysis revealed a significant positive correlation in specific subgroups. In summary, combined exposure increased ONIHL risk and affected health. Machine learning effectively predicted ONIHL, but the study had limitations and needed further research.**

**Keywords** Occupational exposure, Noise and dust, Hearing loss, Machine learning algorithms, Risk factors

Occupational noise and dust are common work-related risk factors and have emerged as significant occupational health hazards. Prolonged exposure to occupational noise can damage the auditory sensory cells of the cochlea, resulting in irreversible loss of inner ear hair cells and subsequent hearing impairment[1]. Noise-induced hearing loss (NIHL) is one of the most frequently reported occupational diseases in numerous studies[2]. In addition to its impact on hearing, noise exposure has been associated with various extra-auditory effects, such as adverse cardiovascular outcomes[3],impaired glucose metabolism[4],increased obesity risks[5].In recent years, research on the correlation between occupational noise exposure and systemic non-auditory effects has gained considerable attention. However, the findings across studies remain inconsistent[6–8].

Similarly, dust is a major pollutant in industrial environments. Long-term exposure to dust can severely damage the respiratory system and contribute to other health problems[9].Despite the prevalence of combined noise and dust exposure in occupational settings, the interaction between these factors and their joint impact on hearing loss has not been fully elucidated. While previous studies have demonstrated that combined exposure to noise and dust increases the risk of hypertension in occupational groups, the potential combined effects on NIHL, glucose metabolism,

obesity, and other health outcomes have not been thoroughly investigated.

With the advancement of technologies, machine learning (ML) algorithms have emerged as.

powerful tools in biomedical research[10]. These algorithms are capable of handling complex datasets with multiple variables, uncovering hidden patterns and relationships that may not be apparent using traditional

[1]Department of Otolaryngology, Hebei Medical University, Shijiazhuang, China. [2]Department of Otolaryngology, Hebei General Hospital, Shijiazhuang, China. [3]Hebei North University, Zhangjiakou, China. [4]Animal Laboratory, The Third Hospital of Hebei Medical University, Shijiazhuang, China. [5]Department of Otolaryngology, The Third Hospital of Hebei Medical University, Shijiazhuang, China. ✉email: 1078652585@qq.com; 36500316@hebmu.edu.cn

statistical methods. Various ML algorithms, such as Logistic Regression, Random Forest Classifier, Decision Tree Classifier, XGB Classifier, AdaBoost Classifier, and Gaussian Naïve Bayes, each have unique strengths in handling different types of data and prediction tasks.

The aim of this study is to leverage machine learning algorithms to investigate the combined effects of noise and dust on hearing loss in occupational populations. By integrating comprehensive datasets, including noise exposure levels, dust concentration, individual characteristics (e.g., age, gender, smoking status), and audiological test results, we aim to construct predictive models capable of accurately assessing the risk of hearing loss in workers exposed to both noise and dust.

This research seeks to provide a deeper understanding of the pathophysiological mechanisms underlying this multifactorial occupational hazard. Furthermore, it aims to offer valuable insights to support the development of effective preventive strategies and occupational health policies, ultimately improving the well-being of workers in high-risk environments.

## Methods

### Study population

A retrospective analysis was conducted on the occupational health examination data of active workers at the Hebei Medical Examination Center, covering the period from January 1, 2017, to December 12, 2023.

All participants were scheduled for examinations in accordance with GBZ188-2014 "Technical Specifications for Occupational Health Surveillance." Grouping was based on exposure to hazardous factors: workers exposed exclusively to noise were categorized into the noise exposure group, while those exposed to both noise and dust (all types of industrial dust) were placed in the noise-dust exposure group.

Dust exposure was defined as exposure to dust in the occupational environment. Although the manuscript does not provide detailed information on dust concentration measurement, the Hebei Medical Examination Center utilized professional dust-measuring instruments to monitor dust levels in the workplace during occupational health assessments. According to relevant occupational health standards, different industries and work scenarios have specific permissible dust exposure limits. Workers in environments where the dust concentration exceeds these limits for a specified duration are classified as dust-exposed.

Noise exposure was defined based on the time-weighted average sound pressure level (TWA) of noise in the workplace. Workers exposed to a TWA of 85 dB(A) or higher over an 8-hour workday were considered to be noise-exposed. This definition aligns with both national and international standards for assessing noise exposure in the context of occupational health.

### Data collection

Inclusion criteria: Workers exposed to both dust and noise. Exclusion criteria: Workers with occupational contraindications for noise-exposed jobs, a history of occupational diseases, or hearing impairments caused by pseudo-hearing loss, exaggerated hearing loss, ototoxic medications (e.g., streptomycin, gentamicin, kanamycin), trauma, infectious diseases (e.g., epidemic cerebrospinal meningitis, mumps, measles), familial hearing loss, Ménière's disease, sudden hearing loss, various middle ear diseases, acoustic neuroma, or auditory neuropathy.

Demographic and clinicopathological data were collected, including gender, age, systolic blood pressure (SBP), diastolic blood pressure (DBP), alcohol consumption, smoking status, electrocardiogram (ECG) findings, fasting blood glucose (FBG), alanine aminotransferase (ALT), body mass index (BMI), and length of service.

*Statistical analyses*

The Chi-square test and Mann-Whitney U test were employed to analyze discrete and continuous variables, respectively. Independent risk factors for occupational noise-induced hearing loss (ONIHL) were identified through multivariable logistic regression analysis using backward stepwise selection. The results were reported as odds ratios (ORs) with 95% confidence intervals (CIs).

Machine learning (ML) algorithms are widely recognized for their superior performance, often surpassing traditional regression methods, especially in predicting outcomes from large and complex datasets[10,11]. In this study, a resampling validation mechanism was utilized to evaluate the performance of each model across multiple training iterations. Key metrics, such as average area under the curve (AUC) scores and their variances, were summarized to assess model performance.

The primary focus was on comparing the overall performance of various models applied to this dataset. The resampling approach ensured consistency in the training samples used across models, enabling more robust and reliable comparisons.

We developed six types of machine learning (ML) algorithms to model the data: Logistic Regression, Random Forest Classifier, Decision Tree Classifier, XGB Classifier, AdaBoost Classifier, and Gaussian Naïve Bayes. In this study, we randomly split the dataset into two groups: the training set (70%) for the development of machine learning models and the validation set (30%) for evaluating model performance. The k-fold training mechanism is designed to select the best model instance from multiple training rounds within a given model type, with the test set serving as an external validation to assess the model's performance. This method is more suitable for optimizing individual models. However, since our goal is to compare multiple models, we chose resampling cross-validation. This method ensures consistency in the training samples selected across multiple model trainings, thus facilitating a better comparison of different models.

To enhance the transparency and reproducibility of the analysis and to effectively prevent overfitting, we employed 5-fold cross-validation as the hyperparameter tuning strategy during model training. Specifically, we divided the dataset into 5 subsets, using 4 subsets for model training and the remaining 1 subset for validation, repeating this process multiple times. Ultimately, we selected the hyperparameter combination that performed the best across all validation sets to ensure the stability and generalizability of the model across different data subsets.

Subsequently, R software was used to further train the ML algorithms for predicting the risk of occupational noise-induced hearing loss (ONIHL) and to evaluate the predictive performance of each classifier on the validation set. Performance metrics included the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and overall accuracy. For model comparison, the closer the ROC curve was to 1, the better the classification performance of the model.

Finally, to explore the relationship between dust exposure and ONIHL in specific populations, stratified analyses were conducted across different subgroups.

All statistical analyses were performed using R (version 4.2.3) and Python (version 3.11.4).

### *Ethical approval and consent to participate*

The study was approved by the Ethics Committee of Hebei General Hospital, and all research was conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants and/or their legal guardians prior to their involvement in the study. A total of 14,145 participants who met the inclusion criteria were included. The Ethics Committee waived the requirement for written informed consent. All patient information was anonymized before analysis to ensure confidentiality.

## Result

### Demographics features

This study included a total of 14,145 workers exposed to occupational noise, among whom 627 were diagnosed with occupational noise-induced hearing loss (ONIHL), while 13,518 were not (Table 1). Significant differences were observed between the two groups in terms of gender, exposure to both noise and dust, systolic and diastolic blood pressure, smoking status, alcohol consumption, fasting blood glucose levels, body mass index (BMI), age, and length of service. However, no significant differences were found in electrocardiogram (ECG) findings or liver function.

Furthermore, the study subjects were stratified based on the presence or absence of combined exposure to noise and dust to investigate its impact on blood pressure, BMI, ECG, fasting blood glucose, and liver function. The results indicated significant differences between the two groups in diastolic blood pressure, BMI, fasting blood glucose levels, and abnormal liver function (Table 2).

### Univariate and multivariate logistic regression analyses of ONIHL

Univariate logistic regression analysis revealed that gender, combined exposure to noise and dust, systolic blood pressure, diastolic blood pressure, smoking, alcohol consumption, BMI, fasting blood glucose, age, and length of service were significantly associated with the occurrence of occupational noise-induced hearing loss (ONIHL) in the overall population ($P < 0.05$). However, abnormal liver function and electrocardiogram (ECG) findings did not show significant differences (Table 3).

Multivariate logistic regression analysis identified male gender (OR = 2.86, 95% CI 1.95–4.22), combined exposure (OR = 11.27, 95% CI 1.06–1.52), elevated diastolic blood pressure (OR = 1.56, 95% CI 1.22–22.01), smoking (OR = 1.69, 95% CI 1.40–2.05), elevated fasting blood glucose (OR = 2.39, 95% CI 1.65–3.47), and age (OR = 1.06, 95% CI 1.05–1.08) as independent positive predictors of ONIHL. In contrast, elevated systolic blood pressure (OR = 0.70, 95% CI 0.51–0.95) was a negative predictor of ONIHL. The multivariate analysis excluded alcohol consumption, BMI, and years of work experience as predictors.

#### *Performance of machine learning algorithms*

Table 4; Fig. 1 present a detailed comparison of the predictive performance of six machine learning (ML) algorithm models on the validation set. The results show that the logistic regression model achieved the highest predictive performance for occupational noise-induced hearing loss (ONIHL), with an ROC of 0.714, sensitivity of 0.700, specificity of 0.621, and accuracy of 0.624. Based on these results, the logistic regression model was selected as the final predictive model.

Compared to Random Forest, Logistic Regression performs better in terms of AUC and sensitivity, indicating a stronger ability to identify positive samples and more effectively reduce the risk of false negatives. However, Logistic Regression has lower specificity, meaning its performance in identifying healthy individuals is less optimal, which may lead to a higher rate of false positives. Additionally, Logistic Regression assumes a linear relationship between the independent and dependent variables, which limits its ability to handle complex nonlinear relationships.

In contrast, the Random Forest model excels in accuracy and specificity, indicating better performance in correctly identifying healthy individuals and overall classification accuracy, thus reducing the risk of false positives. Moreover, Random Forest has strong processing capabilities, enabling it to capture complex nonlinear relationships between features, and it performs robustly in the presence of noisy data and missing values. However, with a lower AUC and poorer sensitivity, it is less effective than Logistic Regression in identifying positive samples.

#### *Relative importance of variables in machine learning algorithms*

In the analysis of permutation importance for ONIHL prediction across various machine learning algorithms (Fig. 2), a clear pattern emerges. While different models yield varying assessments of variable importance, age and smoking consistently rank among the most influential factors, demonstrating strong predictive value. In contrast, the five variables—dust exposure, gender, fasting blood glucose (FBG), systolic blood pressure (SBP),

| Variables | Overall ($n=14145$) | Without ONIHL ($n=13518$) | With ONIHL ($n=627$) | *P*-value |
|---|---|---|---|---|
| **Gender** | | | | $P<0.001$ |
| *Female* | 2401(16.974) | 2371(17.540) | 30(4.785) | |
| *Male* | 11,744(83.026) | 11,147(82.460) | 597(95.215) | |
| **Expouse** | | | | $P=0.002$ |
| *NE* | 4912(34.726) | 4730(34.990) | 182(29.027) | |
| *NCWDE* | 9233(65.274) | 8788(65.010) | 445(70.973) | |
| **SBP** | | | | $P=0.007$ |
| *No* | 12,760(90.209) | 12,214(90.354) | 546(87.081) | |
| *Yes* | 1385(9.791) | 1304(9.646) | 81(12.919) | |
| **DBP** | | | | $P<0.001$ |
| *No* | 12,101(85.550) | 11,617(85.937) | 484(77.193) | |
| *Yes* | 2044(14.450) | 1901(14.063) | 143(22.807) | |
| **Smoke** | | | | $P<0.001$ |
| *No* | 9718(68.703) | 9401(69.544) | 317(50.558) | |
| *Yes* | 4427(31.297) | 4117(30.456) | 310(49.442) | |
| **Alcohol** | | | | $P<0.001$ |
| *No* | 10,443(73.828) | 10,048(74.331) | 395(62.998) | |
| *Yes* | 3702(26.172) | 3470(25.669) | 232(37.002) | |
| **ECG** | | | | $P=0.295$ |
| *No* | 11,273(79.696) | 10,763(79.620) | 510(81.340) | |
| *Yes* | 2872(20.304) | 2755(20.380) | 117(18.660) | |
| **FBG** | | | | $P<0.001$ |
| *No* | 13,879(98.119) | 13,288(98.299) | 591(94.258) | |
| *Yes* | 266(1.881) | 230(1.701) | 36(5.742) | |
| **ALF** | | | | $P=0.832$ |
| *No* | 12,187(86.158) | 11,645(86.144) | 542(86.443) | |
| *Yes* | 1958(13.842) | 1873(13.856) | 85(13.557) | |
| **BMI** | | | | $P=0.019$ |
| *normal* | 5560(39.307) | 5347(39.555) | 213(33.971) | |
| *overweight* | 5495(38.848) | 5233(38.711) | 262(41.786) | |
| *obesity* | 3090(21.845) | 2938(21.734) | 152(24.242) | |
| **Length of service** | 10.000(4.000,19.000) | 10.000(4.000,19.000) | 15.000(5.000,25.000) | $<0.001$ |
| **Age** | 41.000(35.000,49.000) | 41.000(35.000,49.00) | 48.000(40.000,52.000) | $<0.001$ |

**Table 1**. Characteristics of the study population. NE: Noise exposure; NCWDE: Noise combined with dust exposure; SBP: Systolic blood pressure; DBP: Diastolic blood pressure; FBG:Fasting blood glucose; ALF:Abnormal liver function.

and diastolic blood pressure (DBP)—exhibit lower predictive contributions and show considerable fluctuations in their importance rankings across different models.

Age and smoking are well-established risk factors for ONIHL, with extensive epidemiological evidence supporting their strong association with hearing loss. Consequently, their importance remains relatively stable across different machine learning models, resulting in consistent rankings. In contrast, the impact of other variables on ONIHL may be more indirect or context-dependent, with underlying mechanisms that are not yet fully understood. Furthermore, differences in feature selection strategies, weight assignment, and the handling of variable interactions across machine learning models may contribute to the observed inconsistencies in their importance rankings.

Subgroup analysis  Subgroup analysis (Fig. 3) revealed a particularly significant positive correlation between combined exposure to noise and dust and the prevalence of occupational noise-induced hearing loss (ONIHL) in males, individuals aged 30–39 years, those with less than 10 years of work experience, individuals without hypertension, those with a normal BMI, normal fasting blood glucose, normal liver function, and a normal electrocardiogram.

## Discussion

The present study aimed to explore the combined impact of noise and dust exposure on hearing loss and extra-auditory effects, such as cardiovascular health, glucose metabolism, and obesity risks, in occupationally exposed populations. The study also sought to identify associated risk factors using machine learning algorithms. We found that combined exposure to noise and dust significantly increases the risk of noise-induced hearing loss

| Variables | Overall | Without Dust exposure (*n* = 4912) | With Dust exposure (*n* = 9233) | *P*-value |
|---|---|---|---|---|
| **SBP** | | | | |
| <140 | 12,760(90.209) | 4400(89.577) | 8360(90.545) | *P* = 0.079 |
| ≥ 140&<160 | 1172(8.286) | 438(8.917) | 734(7.950) | |
| ≥ 160&<180 | 203(1.435) | 73(1.486) | 130(1.408) | |
| ≥ 180 | 10(0.071) | 1(0.020) | 9(0.097) | |
| **DBP** | | | | |
| <90 | 12,101(85.550) | 4136(84.202) | 7965(86.267) | *P* = 0.005 |
| ≥ 90&<100 | 1356(9.586) | 520(10.586) | 836(9.054) | |
| ≥ 100&<110 | 550(3.888) | 211(4.296) | 339(3.672) | |
| ≥ 110 | 138(0.976) | 45(0.916) | 93(1.007) | |
| **BMI** | | | | |
| <24 | 5560(39.307) | 1972(40.147) | 3588(38.861) | *P* < 0.001 |
| ≥ 24&<28 | 5495(38.848) | 1795(36.543) | 3700(40.074) | |
| ≥ 28 | 3090(21.845) | 1145(23.310) | 1945(21.066) | |
| **ECG** | | | | |
| No | 11,273(79.696) | 3907(79.540) | 7366(79.779) | *P* = 0.736 |
| Yes | 2872(20.304) | 1005(20.460) | 1867(20.221) | |
| **FBG** | | | | |
| No | 13,879(98.119) | 4848(98.697) | 9031(97.812) | *P* < 0.001 |
| Yes | 266(1.881) | 64(1.303) | 202(2.188) | |
| **ALF** | | | | |
| No | 12,187(86.158) | 4136(84.202) | 8051(87.198) | *P* < 0.001 |
| Yes | 1958(13.842) | 776(15.798) | 1182(12.802) | |

**Table 2**. The influence of combined dust exposure on blood pressure, BMI, etc.

| Variables | Univariable analysis | | Multivariable analysis | |
|---|---|---|---|---|
| | OR 95% CI | *P*-value | OR 95% CI | *P*-value |
| **Gender** | 4.23 (2.93–6.12) | *P* < 0.001 | 2.86 (1.95–4.22) | *P* < 0.001 |
| **Expouse** | 1.32 (1.10–1.57) | *P* = 0.002 | 1.27 (1.06–1.52) | *P* = 0.011 |
| **SBP** | 1.39 (1.09–1.77) | *P* = 0.007 | 0.70 (0.51–0.95) | *P* = 0.021 |
| **DBP** | 1.81 (1.49–2.19) | *P* < 0.001 | 1.56 (1.22–2.01) | *P* < 0.001 |
| **Smoke** | 2.23 (1.90–2.62) | *P* < 0.001 | 1.69 (1.40–2.05) | *P* < 0.001 |
| **Alcohol** | 1.70 (1.44–2.01) | *P* < 0.001 | 1.08 (0.89–1.31) | *P* = 0.455 |
| **BMI** | | | | |
| *normal* | *Reference* | | *Reference* | |
| *overweight* | 1.26 (1.04–1.51) | *P* = 0.015 | 0.88 (0.73–1.07) | *P* = 0.209 |
| *obesity* | 1.30 (1.05–1.61) | *P* = 0.016 | 0.96 (0.77–1.21) | *P* = 0.744 |
| **FBG** | 3.52 (2.45–5.05) | *P* < 0.001 | 2.39 (1.65–3.47) | *P* < 0.001 |
| **ALF** | 0.98 (0.77–1.23) | *P* = 0.832 | | |
| **ECG** | 0.90 (0.73–1.10) | *P* = 0.295 | | |
| **Age** | 1.06 (1.05–1.07) | *P* < 0.001 | 1.06 (1.05–1.08) | *P* < 0.001 |
| **Length of service** | 1.04 (1.03–1.04) | *P* < 0.001 | 0.99 (0.98–1.00) | *P* = 0.062 |

**Table 3**. Univariate and multivariate logistic regression analysis of variables in predicting ONIHL.

(NIHL), as well as diastolic blood pressure, BMI, fasting blood glucose, and abnormal liver function, compared to exposure to noise alone.

Furthermore, male gender, age, combined exposure, elevated diastolic blood pressure, elevated fasting blood glucose, and smoking were identified as independent positive predictors of ONIHL. The results of multivariate logistic regression were validated using machine learning algorithms, with the logistic regression model confirmed as the optimal predictive model. This finding was consistent with the multivariate analysis. Additionally, men without underlying diseases, aged 30–39 years, and with less than 10 years of work experience, were found to be at higher risk of developing NIHL.

Several cross-sectional studies have investigated the relationship between noise exposure, age, and hearing loss. Somma et al.[12] observed 5 dB and 20 dB hearing losses in younger and older cement workers (exposed to

| Methods | AUC | Sensitivity | Specificity | Accuracy |
|---|---|---|---|---|
| Logistic | 0.714 | 0.700 | 0.621 | 0.624 |
| RandomForest | 0.638 | 0.490 | 0.726 | 0.715 |
| DecisionTree | 0.625 | 0.528 | 0.705 | 0.697 |
| XGBoost | 0.693 | 0.571 | 0.708 | 0.702 |
| AdaBoost | 0.709 | 0.693 | 0.616 | 0.619 |
| GNB | 0.695 | 0.752 | 0.545 | 0.554 |

**Table 4**. Predictive performance comparison of the six types of machine learning algorithms in the validation sets.

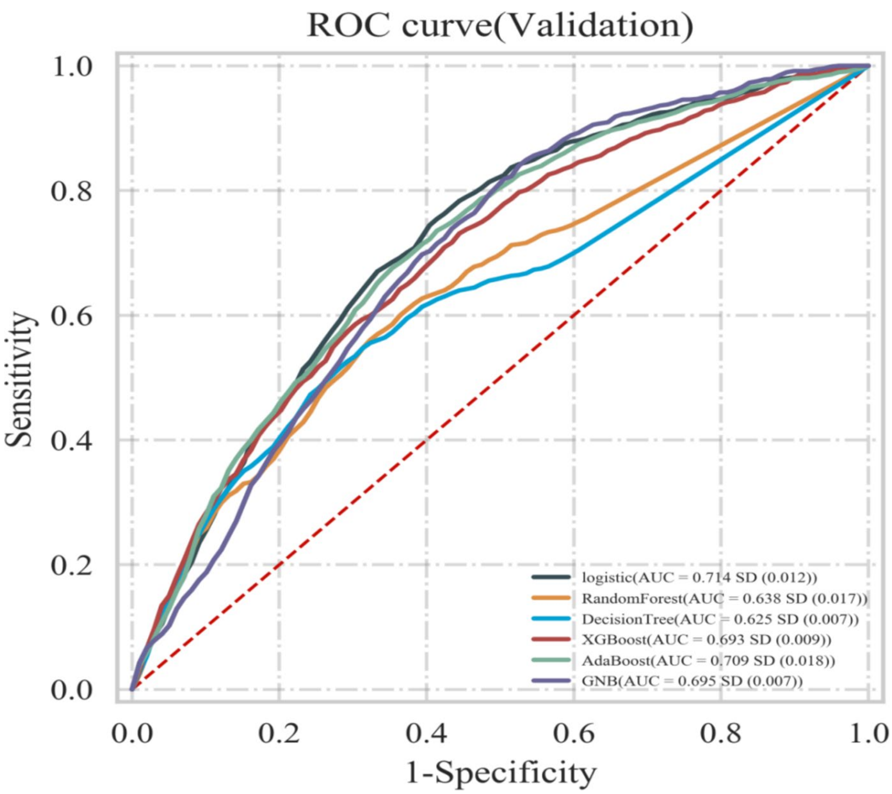


**Fig. 1**. ROC Curve Analysis of Machine Learning Algorithms for Predicting ONIHL Using Gender, Exposure, SBP, DBP, Smoking Status, FBG, and Age.

> 85 dB daily noise) compared to controls. Golmohammadi et al.[13]found that hearing loss increased with age and years of exposure in 1,062 tractor workers (Leq > 85 dB(A)), with specific regression coefficients. Hederstierna and Rosenhall[14] reported an additive relationship between noise-induced hearing loss (NIHL) and aging in individuals aged 70–75 years.

Overall, the literature supports an additive model of NIHL and aging, which is consistent with our findings. Specifically, our study demonstrated that the detection rate of NIHL increases with both years of exposure and age in workers exposed to noise.

The detection rate of noise-induced hearing loss (NIHL) was higher in males than in females, and this difference was statistically significant, which is consistent with previous studies[15,16].This may be attributed to differences in hormone levels, as well as the higher prevalence of harmful habits such as smoking and alcohol consumption in males compared to female[17],Our results also align with these findings, showing higher detection rates of NIHL in populations with smoking and alcohol consumption habits.

Additionally, our study indicated that individuals with hypertension, high fasting blood glucose, and a high body mass index (BMI), particularly those who are obese, are more vulnerable to NIHL. The underlying mechanisms are as follows: In cases of hyperglycemia, persistent high blood sugar levels damage the inner ear by thickening the microvessel basement membranes, leading to microvascular lesions, impaired nerve fiber conduction, and disrupted metabolism, all of which ultimately impair hearing[18].

Hypertension similarly disrupts cochlear microcirculation, increasing the risk of hemorrhage and subsequent reperfusion injury. Damaged hair cells and reduced blood flow further lower noise tolerance, enhancing the risk

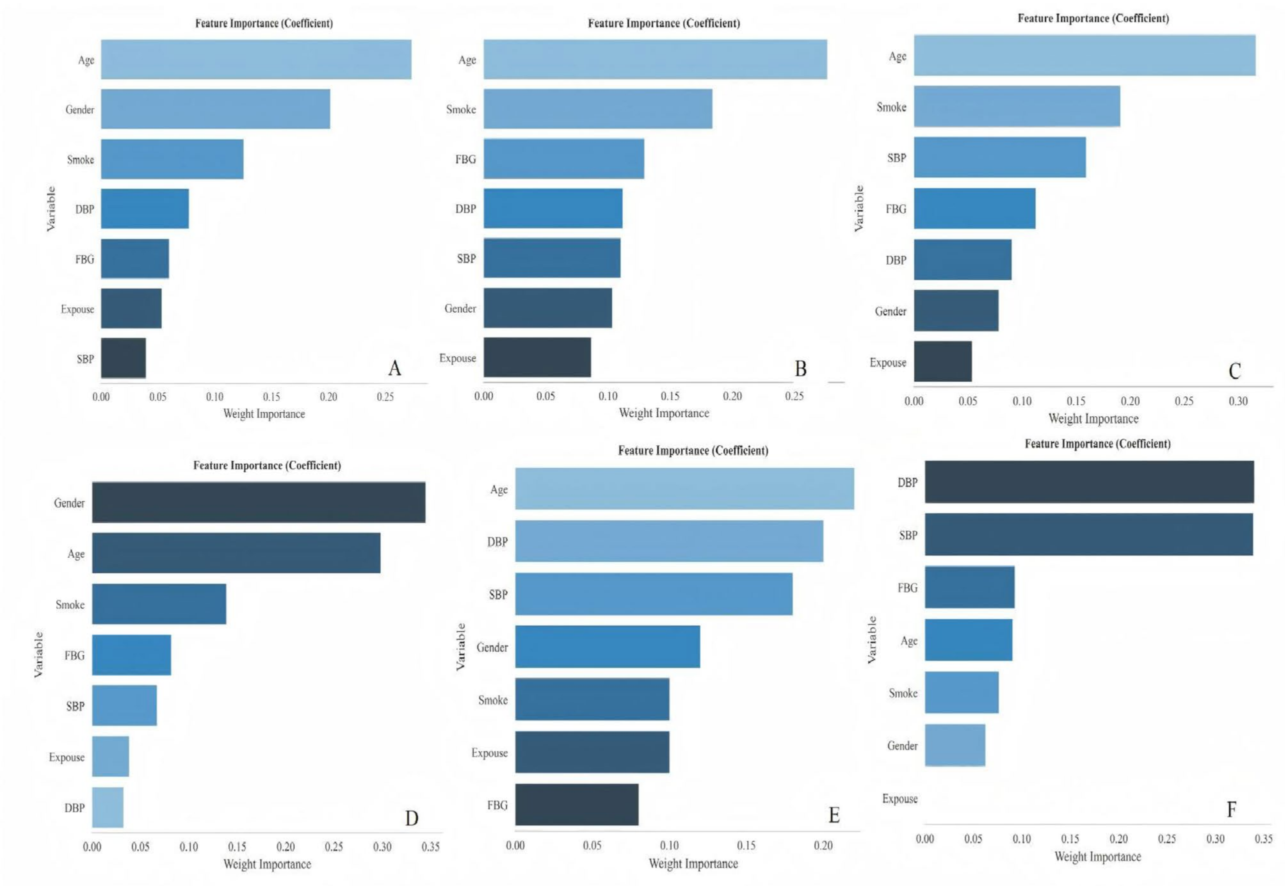


**Fig. 2**. SHAP plots illustrating the contribution of each input variable to ONIHL prediction across various machine learning algorithms: (**A**) Logistic Regression, (**B**) Random Forest, (**C**) Decision Tree, (**D**) XGBoost, (**E**) AdaBoost, and (**F**) Gaussian Naïve Bayes (GNB).

of hearing loss. Systolic and diastolic blood pressure have different physiological mechanisms. Therefore, in this study, systolic and diastolic pressures were analyzed separately. Systolic pressure reflects the volume of blood ejected by the heart during contraction and the tension in the arterial walls, while diastolic pressure reflects the pressure in the arteries when the heart is relaxed. Their effects on blood vessels and organs are distinct, so analyzing them separately provides a clearer understanding of their different roles in health. Systolic pressure likely has a more significant impact on hemodynamics and cochlear blood supply, whereas diastolic pressure may play a more crucial role in long-term vascular health. Additionally, clinical treatment and interventions typically address systolic and diastolic pressures with different strategies. Thus, separate analysis allows for a more accurate assessment of their independent effects and helps avoid confounding their influence.

Obesity also contributes to inner ear damage. Vascularly, it increases capillary pressure and promotes atherosclerosis, while metabolically, it reduces adiponectin levels and induces oxidative stress. These combined effects contribute to damage in the inner ear[19].

Furthermore, a novel finding in our study is that combined exposure to noise and dust significantly increases the risk of NIHL in workers compared to single noise exposure. This suggests a synergistic effect of noise and dust on the development of NIHL.

Previous research has shown that noise exposure acutely affects the autonomic nervous and endocrine systems[3],, influencing blood pressure (BP) levels and altering blood lipid and glucose concentrations, which are key factors in the development of cardiovascular disease[18].Occupational noise exposure has also increasingly been identified as a risk factor for obesity. The potential mechanisms underlying this association may involve noise exposure inducing dysregulation of the endocrine system through activation of the hypothalamus-pituitary-adrenal axis and elevated cortisol levels. This, in turn, can alter metabolism, promote central fat deposition, and contribute to obesity[5,19].Additionally, some studies indicate that long-term dust exposure leads to lung fibrosis and the entry of dust particles into the bloodstream, causing stress reactions in other systems. This can lead to the development of pulmonary hypertension, accelerate atherosclerosis, and further affect the function of cardiovascular and other systems[23]. In our study, we grouped participants based on the presence or absence of combined exposure and found significant differences between the two groups in terms of diastolic blood pressure (DBP), BMI (especially obesity), fasting blood glucose, and liver function. However, we did not observe significant changes in the prevalence of DBP or abnormal electrocardiogram (ECG) in exposed

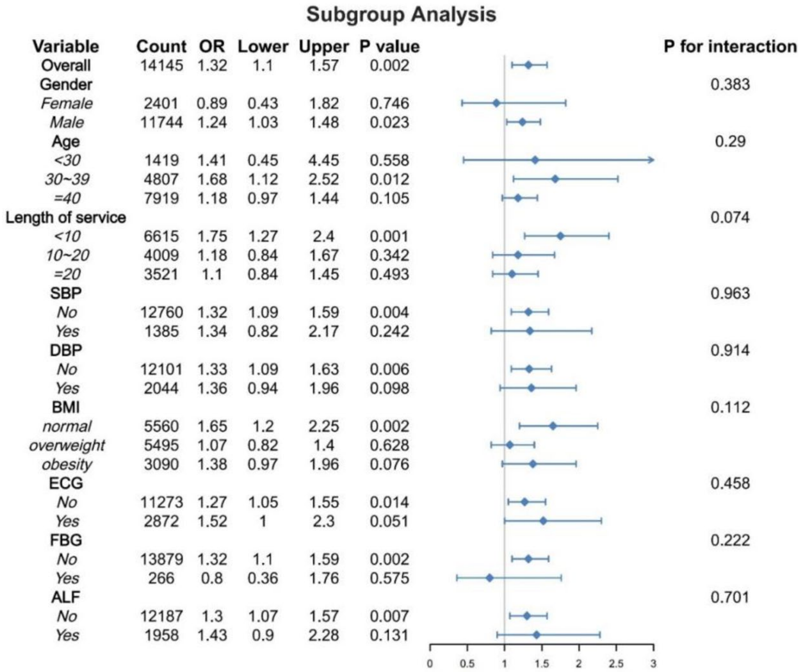


**Fig. 3**. Stratified Forest plot of the association between dust exposure and risk of ONIHL.

participants, although there was an increase in systolic blood pressure (SBP) and abnormal liver function.Our findings are largely consistent with those from a controlled exposure study, which showed an increase in DBP but no change in SBP or heart rate in 18 healthy human volunteers exposed to occupational noise at 95 dBA for 20 min[20]Additionally, exposure to noise exacerbated liver damage[21] .

Our results demonstrated that several factors were significantly associated with the occurrence of noise-induced inner ear hearing loss (ONIHL). Male gender emerged as a strong independent predictor, which is consistent with previous research suggesting that males may be more susceptible to noise-induced hearing loss, possibly due to differences in occupational distribution and exposure levels between genders[2]. Combined exposure to noise and dust was found to have a substantial impact, with an odds ratio of 11.27. This indicates a synergistic effect of these two pollutants on the auditory system, underscoring the importance of considering their combined exposure in occupational health assessments.

Elevated diastolic blood pressure was positively associated with ONIHL, aligning with the growing body of evidence linking cardiovascular health and hearing loss. The underlying mechanism may involve vascular changes in the inner ear, as increased blood pressure could affect the delicate blood supply to the cochlea[3]. Smoking and elevated fasting blood glucose were also identified as independent risk factors. Smoking can cause vasoconstriction and reduce oxygen supply to the cochlea, while hyperglycemia may lead to oxidative stress and damage to the auditory nerve fibers[4,5]. Interestingly, systolic blood pressure was found to be a negative predictor in the multivariate analysis. Consider the possible causes, moderate increases in systolic pressure may improve cochlear blood supply, reducing inner ear damage caused by hypoxia or abnormal blood flow. The labyrinth, which supplies the inner ear, is a terminal artery, making it susceptible to the influence of overall circulation. Stable blood supply ensures its normal function and prevents ischemia-induced hearing loss. Furthermore, elevated systolic pressure often prompts individuals to pay more attention to cardiovascular health, leading to more proactive health management, which may indirectly contribute to hearing protection and reduce the risk of ONIHL. Further research is needed to clarify this relationship.

The application of machine learning (ML) algorithms in this study demonstrated their potential superiority over traditional regression methods in handling the complex relationships within our dataset, particularly when dealing with a large number of samples. The resampling validation mechanism employed was essential in providing a comprehensive assessment of the models' performance across multiple training runs. We developed six types of ML algorithms to model our data: Logistic Regression, Random Forest Classifier, Decision Tree Classifier, XGB Classifier, AdaBoost Classifier, and Gaussian Naive Bayes (NB). After optimizing, adjusting, and training the ML models, we found that the logistic model achieved the highest predictive performance for ONIHL. The results of multivariate logistic regression were further validated by the machine learning algorithms, confirming that logistic regression was the best model and consistent with the prior results of multivariate logistic regression.

To explore the relationship between dust exposure and ONIHL in specific populations, stratified analyses were conducted across different subgroups. The results revealed that men without pre-existing conditions, aged 30–39 years, and with less than 10 years of work experience are at an elevated risk of developing ONIHL. A possible explanation is that certain subgroups may be more vulnerable to the combined effects of noise and dust, and may lack the protective effects of estrogen. Additionally, these individuals are more likely to be in occupations with high noise exposure. This suggests the need for targeted preventive strategies within these populations. For instance, in male-dominated occupations with high noise and dust exposure, stricter hearing protection measures and regular health monitoring should be implemented. Younger workers and those with less work experience may require additional training on the importance of occupational safety and the potential risks of exposure.

The AUC value of 0.714 indicates a moderate discriminatory capacity of the model, although there is room for improvement. Future research could explore more advanced machine learning techniques or incorporate additional relevant variables to enhance predictive precision. The comparison of different algorithm performances provides valuable insights, assisting researchers in selecting appropriate models based on data characteristics and research objectives.

The results of this study have significant implications for occupational health policies and practices. Employers must recognize the substantial risks associated with combined noise and dust exposure and take proactive measures to mitigate these risks. Engineering controls, the provision of personal protective equipment, and proper workplace ventilation are crucial components of these efforts. Regular health check-ups should include comprehensive assessments of hearing, cardiovascular health, and metabolic markers to identify potential health issues at an early stage. Additionally, health promotion initiatives should encourage workers to adopt healthier lifestyles, such as quitting smoking, maintaining a balanced diet, and engaging in regular physical activity, which could potentially counteract the harmful effects of occupational exposures.

## Limitations

Despite the valuable findings of this study, several limitations should be acknowledged. The study was based on a retrospective review of occupational health examination data, which may be subject to recall bias and selection bias. Future studies could consider prospective cohort designs to more robustly establish causal relationships. Additionally, although we included a relatively comprehensive set of variables, there may be other unmeasured factors that could influence the results. For instance, genetic factors related to hearing loss susceptibility and more detailed information on specific types of dust and noise exposure could be further explored. Moreover, the generalizability of our findings may be limited to the specific population studied, and additional research in different regions and industries is needed to confirm the consistency of the results.

## Conclusion

In conclusion, this study highlights the combined impact of noise and dust on hearing loss in occupational settings and identifies key risk factors. Machine learning algorithms offer a promising approach for predicting ONIHL, though further refinement is necessary. Combined exposure increases the risk of ONIHL and is associated with elevated diastolic blood pressure, BMI, fasting blood glucose, and liver dysfunction, compared to exposure to noise alone. Male gender, age, combined exposure, elevated diastolic blood pressure, elevated fasting blood glucose, and smoking were identified as significant predictors of ONIHL. The algorithms confirmed the validity of multivariate logistic regression, aligning with its results and supporting its optimality. Notably, men aged 30–39, without pre-existing conditions, and with less than 10 years of work experience, are at a higher risk of ONIHL. These findings have important implications for occupational health and will guide future research.

## Data availability

The dataset of the present study is available upon reasonable request from the corresponding author.

Received: 14 January 2025; Accepted: 11 March 2025
Published online: 17 March 2025

## References

1. McGill, T. J. & Schuknecht, H. F. Human cochlear changes in noise induced hearing loss. *Laryngoscope* **86**, 1293–1302 (1976).
2. Lie, A. et al. Occupational noise exposure and hearing: a systematic review. *Int. Arch. Occup. Environ. Health*. **89**, 351–372 (2016).
3. Lusk, S. L., Gillespie, B., Hagerty, B. M. & Ziemba, R. A. Acute effects of noise on blood pressure and heart rate. *Arch. Environ. Health*. **59**, 392–399 (2004).
4. Skogstad, M. et al. Systematic review of the cardiovascular effects of occupational noise. *OCCMED* **66**, 10–16 (2016).
5. An, R., Wang, J., Ashrafi, S. A., Yang, Y. & Guan, C. Chronic noise exposure and adiposity: A systematic review and Meta-analysis. *Am. J. Prev. Med.* **55**, 403–411 (2018).
6. Chang, T. Y. et al. Occupational noise exposure and incident hypertension in men: a prospective cohort study. *Am. J. Epidemiol.* **177**, 818–825 (2013).
7. Liu, T. et al. [Health hazards and hearing loss risk assessment of workers exposed to noise in an automobile manufacturing enterprise]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi.* **40**, 434–438 (2022).
8. Kerns, E., Masterson, E. A., Themann, C. L. & Calvert, G. M. Cardiovascular conditions, hearing difficulty, and occupational noise exposure within US industries and occupations. *Am. J. Industrial Med.* **61**, 477–491 (2018).
9. Guo, J. et al. Long-Term exposure to dust and noise can increase the risk of Hypertension—A population based, Cross-sectional study. *J. Occup. Environ. Med.* **65**, e128–e133 (2023).
10. Ngiam, K. Y. & Khor, I. W. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* **20**, e262–e273 (2019).
11. Hulsen, T. et al. From big data to precision medicine. *Front. Med.* **6**, 34 (2019).
12. Somma, G. et al. Extended high-frequency audiometry and noise induced hearing loss in cement workers. *Am. J. Industrial Med.* **51**, 452–462 (2008).
13. Study of occupational. noise - induced hearing loss in a tractor manufacturing plant | Semantic Scholar. https://www.semanticscholar.org/paper/Study-of-occupational-noise-induced-hearing-loss-in-Golmohammadi-sardrudi/dabcbca4b4a7b8ee57900dc9bb6379896be58fdc
14. Hederstierna, C. & Rosenhall, U. Age-related hearing decline in individuals with and without occupational noise exposure. *Noise Health*. **18**, 21–25 (2016).

15. Noise-induced hearing injuries, active component & Armed Forces, U. S. 2007–2010.,MSMR - X-MOL. https://www.x-mol.com/paper/1212923426866274308?adv
16. Nash, S. D. et al. The prevalence of hearing impairment and associated risk factors: the beaver dam offspring study. *Arch. Otolaryngol. Head Neck Surg.* **137**, 432 (2011).
17. Hsieh, C. J. et al. Arctigenin, a dietary phytoestrogen, induces apoptosis of Estrogen receptor-negative breast cancer cells through the ROS/p38 MAPK pathway and epigenetic regulation. *Free Radic Biol. Med.* **67**, 159–170 (2014).
18. Li, W. et al. Fasting serum total bile acid level is associated with coronary artery disease, myocardial infarction and severity of coronary lesions. *Atherosclerosis* **292**, 193–200 (2020).
19. Rosmond, R. et al. A 5-year follow-up study of disease incidence in men with an abnormal hormone pattern. *J. Intern. Med.* **254**, 386–390 (2003).
20. Andrén, L., Hansson, L., Björkman, M. & Jonsson, A. Noise as a contributory factor in the development of elevated arterial pressure: A study of the mechanisms by which noise May Raise blood pressure in man. *J. Intern. Med.* **207**, 493–498 (1980).
21. Belji Kangarlou, M. et al. Noise and silver nanoparticles induce hepatotoxicity via CYP450/NF-Kappa B 2 and p53 signaling pathways in a rat model. *Toxicol. Ind. Health*. **40**, 206–219 (2024).

## Author contributions

Y.Li. and Y.Qu. hypothesized this work. Y.Li. designed the study.Y.Li.,X.Sun. performed data collection, data statistical analysis and drafting of the manuscript; X.Sun. prepared the figures.Y.Qu, S.Yang and Y. Zhai. contributed to the data collection. All authors reviewed the manuscript.

## Funding

This study was supported by the The central government guides local science and technology development fund projects(236Z7730G).

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Y.L. or Y.Q.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.