Preprints.org

Article

Elucidating the Causal Link Between Hearing and Chronic Kidney Disease Through Mendelian Randomization and Development of Predictive Machine Learning Models Based on NHANES Data

Yanting Yang , Yuqing Li , Zhongquan Huang , Rongrong Lin , Zhuoni Chen , Ning Zhang , Nan Zheng [*] ,
Jiean Xu [*] , Yaqi Zhou [*]

Article

Elucidating the Causal Link Between Hearing and Chronic Kidney Disease Through Mendelian Randomization and Development of Predictive Machine Learning Models Based on NHANES Data

Yanting Yang [1,2,†], Yuqing Li [3,†], Zhongquan Huang [3,†], Rongrong Lin [3], Zhuoni Chen [3], Ning Zhang [2], Nan Zheng [4,*] Jiean Xu [5,*] and Yaqi Zhou [1,*]

[1]   State Key Laboratory of Chemical Oncogenomics, Key Laboratory of Chemical Genomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen 518055, China; 2813134542@qq.com (Y.Y.); zhou_yuanie@163.com (Y.Z.)

[2]   School of Basic Medical Sciences, Guangzhou University of Chinese Medicine, Guangzhou 510006, China; 2813134542@qq.com (Y.Y.); 2017101060@stu.gzucm.edu.cn (N.Zhang.)

[3]   School of Pharmaceutical Science, Guangzhou University of Chinese Medicine, Guangzhou 510006, China; 20211111033@stu.gzucm.edu.cn (Y.L.); kevin1737@163.com (Z.H.); 979244995@qq.com (R.L.); 3461880238@qq.com (Z.C.)

[4]   Center for Translational Research, Shenzhen Bay Laboratory, Shenzhen 518000, China; zhengnan@szbl.ac.cn (N.Zheng.)

[5]   Research Centre of Basic Integrative Medicine, Department of Physiology, School of Basic Medical Sciences, Guangzhou University of Chinese Medicine, University Town, Guangzhou 510006, China; xujiean@gzucm.edu.cn (J.X.)

*   Correspondence: zhengnan@szbl.ac.cn (N.Zheng.); xujiean@gzucm.edu.cn (J.X.); zhou_yuanie@163.com (Y.Z.)

†   These authors contributed equally to this work and share first authorship.

Abstract: Chronic kidney disease (CKD) is a growing public health issue with significant morbidity and mortality rates. Traditional diagnostics, with about 70% accuracy, often delay CKD detection, highlighting the need for more efficient methods. Recent studies indicate a potential link between CKD and auditory health, yet this relationship remains underexplored due to methodological challenges and the complexity of establishing causality. Our research utilizes data spanning from 2000 to 2020 from the National Health and Nutrition Examination Survey (NHANES), covering 12,392 participants, including 2,060 diagnosed with CKD. Through meticulous analysis employing logistic regression and Mendelian randomization, we have unearthed novel insights into the bi-directional associations between hearing impairment and CKD. Furthermore, we developed and validated a machine learning model that surpasses traditional diagnostic approaches in terms of accuracy and predictive power. These findings highlight the innovative integration of auditory examinations with demographic data to enhance CKD detection. Our approach demonstrates the potential of machine learning in transforming diagnostic methodologies, thus offering a significant advancement in the field of nephrology and public health.

Keywords: chronic kidney disease; hearing; mendelian randomization; machine learning models; predictions

1. Introduction

Chronic Kidney Disease (CKD) represents a significant global public health concern. The prevalence of CKD is on an upward trajectory, exacerbated by demographic aging, escalating environmental pollution, and shifts in lifestyle patterns. The Global Burden of Disease Consortium predicts that CKD will become the top five leading diseases leading to life lost by 2040 [1]. Beyond

its direct impact on renal function, CKD substantially elevates the risk of myriad complications; notably, cardiovascular mortality rates in CKD patients are estimated to be 2 to 3 times higher than in the general population [2], with hypertension prevalence ranging from 65% to 95% [3]. Furthermore, diabetes emerges as the etiology for half of all end-stage renal disease (ESRD) cases [4]. The current approach to the diagnosis of CKD predominantly encompasses traditional methodologies, including but not limited to blood and urine analyses, as well as imaging procedures (e.g., ultrasound, computed tomography [CT], magnetic resonance imaging [MRI]), renal biopsy, and histopathological examination [5]. Nevertheless, it is observed that the precision of conventional diagnostic practices for CKD typically hovers around the 70% mark [6]. The indistinct early symptomatology of CKD significantly enhances the risk of misdiagnosis or delayed diagnosis, underscoring the imperative for more streamlined, rapid, and efficacious diagnostic approaches [7].

The ancient Chinese medical book, *Huangdi Neijing* (Inner Canon of Yellow Emperor), posits that "the kidney manifests in the ear," suggesting a potential intimate link between ear health and kidney function. The nexus between CKD and auditory anomalies has garnered scholarly attention since the onset of the 20th century, with observational studies delving into the link between CKD and auditory function. These studies report an increased prevalence of sensorineural hearing loss (SNHL) among patients with CKD, a type of hearing loss directly related to damage within the inner ear or auditory nerve [8–11]. However, research assessing the correlation between CKD and hearing loss has primarily focused on hearing loss at different frequencies, with insufficient consideration of other ear-related factors such as tympanic chamber measurements and otoscopic examinations. The variability in study outcomes may be attributed to differences in hearing assessment methodologies (pure-tone audiometry, threshold determination, surveys, etc.) and the influence of confounding factors [12,13]. Current research predominantly scrutinizes the correlation between CKD and auditory impairment without a unified stance on the sequential development of CKD and hearing loss or establishing a definitive causal relationship. To delve deeper into this domain, it is imperative to employ Mendelian randomization analysis to investigate the causal relationships between CKD-related factors and hearing disorders. We investigate the causal relationships between other specified disorders of kidney and ureter and age-related hearing impairment (MTAG), as well as the causality between kidney injury molecule levels and MTAG. Molecular markers of kidney injury have been identified as potential biomarkers for the diagnosis and prognosis of CKD, predicting the presence of CKD [14], while other specified disorders of kidney and ureter are classified as subtypes of chronic kidney disease [15].

In the realm of epidemiological research, Mendelian randomization stands as a formidable methodology, wherein genetic variability is leveraged to probe the causal linkage between risk elements and specific health conditions [16]. The presence of confounders significantly impedes the process of causal deduction within these studies. Conversely, genetic variants utilized in Mendelian randomization investigations adhere to the principle of the stochastic allocation of alleles to progeny, mirroring the methodology intrinsic to randomized controlled trials [17]. This technique effectively mitigates the influence of confounding variables and the issue of reverse causality that are prevalent in observational research, while also addressing concerns related to the representativeness and applicability seen in randomized controlled experiments [18]. Consequently, the Mendelian randomization methodology was employed in this investigation to assess the association between factors related to chronic kidney disease and hearing impairment.

The investigation delved into the putative link between otologic indicators and CKD, further scrutinizing the causative associations between subtypes of CKD and auditory impairment. With the advent of information technology, an extensive compilation of clinical data has been amassed and analyzed, furnishing novel insights for the prognostication and diagnosis of ailments. Notably, the deployment of machine learning algorithms has manifested substantial benefits in enhancing diagnostic precision and forecasting disease susceptibility through the adept handling of voluminous linear and nonlinear datasets. Within the domain of CKD diagnosis, the precision of machine learning models has been documented to range between 80 and 90%, markedly surpassing that of conventional diagnostic techniques. Presently, an array of machine learning models, leveraging

algorithms such as support vector machines, random forests, and artificial neural networks, are employed in the prediction of CKD. The datasets for these models are typically sourced from IBM Explorys, hospital records, and other repositories [19–23].

This investigation represents the inaugural effort to assess the prognostic utility of otologic indicators for CKD. Employing multiple machine learning (ML) methodologies sourced from NHANES, which encompassed 12,392 participants, a novel predictive model for early detection of CKD based on auricular metrics was developed and validated. The predictive efficacy of this model was juxtaposed with that of pre-existing models to underscore the enhanced accuracy afforded by the ML algorithm, grounded in a data-driven otologic examination. It is anticipated that the fusion of TCM principles with contemporary medical technologies will pave the way for novel strategies and methodologies in the early detection of CKD, ameliorating patient outcomes, alleviating suffering, and enhancing quality of life.

2. Materials and Methods

2.1. Data Sources

The NHANES, conducted by the United States National Health Statistics Service (NCHS), represents a pivotal cross-sectional survey designed to amass data regarding the health and nutritional status of the U.S. populace. This endeavor facilitates the acquisition of critical insights into the health conditions prevalent within the nation, thereby underpinning the formulation of informed policies by governmental entities. Endorsement of the NHANES survey protocol by the NCHS Research Ethics Review Board was secured, accompanied by the provision of informed written consent from all participants. Accessibility to the entire database (https://wwwn.cdc.gov/nchs/nhanes/Default.aspx) has rendered the requirement for an ethics review in our research exempt [24].

The cohort for this investigation was derived from NHANES dataset, covering the period from 2000 to 2020. Inclusion criteria mandated a diagnosis of CKD and the availability of comprehensive otology examination data. Individuals lacking complete baseline demographic information or missing audiological test results were systematically excluded (**Figure 1**). Following rigorous selection criteria focused on extensive exposure assessment, a total of 12,392 participants were incorporated into the final analysis. This meticulous approach ensured the integrity and relevance of the data, which underpin further conclusions.

2.2. Statistical Analysis

In this study, a complex sampling design along with sampling weights, in alignment with the guidelines for analysis set forth by NHANES, was meticulously applied to ensure that the findings accurately reflect the demographic composition of the U.S. populace. This methodology facilitated precise estimations of the prevalence of CKD and its associated health metrics within the general population. For the initial data assessment, the Cochran-Mantel-Haenszel test was employed to elucidate disparities in the distribution of categorical variables, namely sex, race, and citizenship status, between individuals diagnosed with CKD and those without, while accommodating for stratification effects. Moreover, differences in continuous variables such as age and the poverty index between these groups were scrutinized using the *t*-test.

Subsequently, logistic regression was applied to identify auditory test measures associated with CKD. Multivariate logistic regression analysis was employed to ascertain the factors correlated with CKD. This methodology facilitated the discovery of specific variables significantly associated with an increased risk of CKD, allowing for the isolation of these effects from other confounding factors. Through this analytical approach, factors that maintained a significant association with CKD risk were identified, with the provision of corresponding odds ratios (ORs) and their 95% confidence intervals (CIs), thereby enhancing the accuracy and relevance of the study's outcomes.

Furthermore, Pearson's correlation analysis was applied across 37 variables within the CKD dataset to quantify the extent of their influence on the condition's presence or absence, with the

correlation coefficients indicating the degree of association. This comprehensive analysis underscored the intricate interplay between various factors and CKD, thereby contributing to a deeper understanding of its epidemiology.
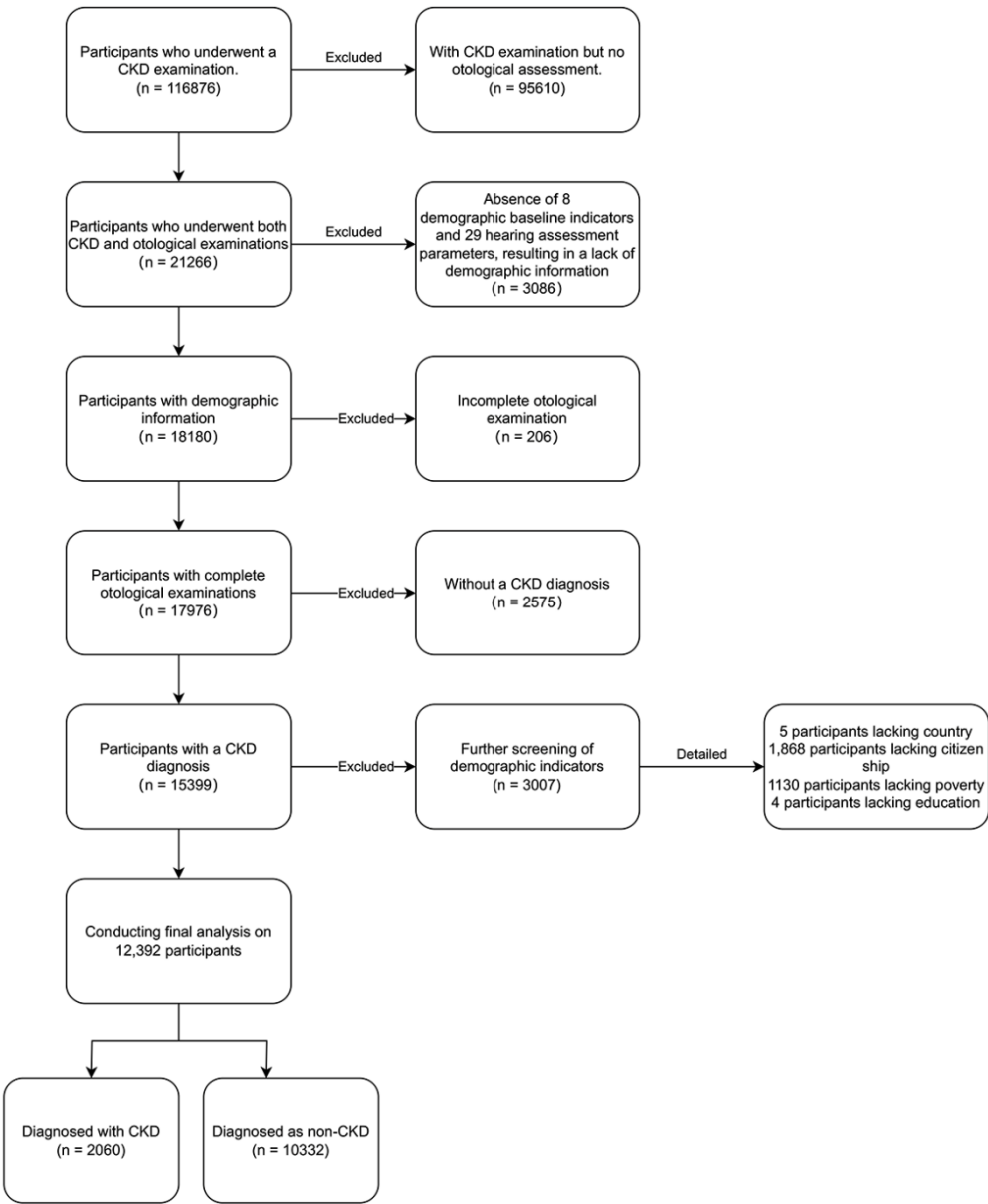
Figure 1. Participant selection flowchart.

2.3. Mendelian Randomization Analysis

2.3.1. Data Sources

Genome-wide association studies (GWAS) have collated data on hearing impairment, various kidney and ureteral diseases, and kidney injury at the molecular level through the MRC IEU OpenGWAS data infrastructure [25]. This research predominantly utilized exposure and outcome data derived from the UK Biobank (http://www.nealelab.is/uk-biobank) and FinnGen Biobank (https://www.finngen.fi/fi). Exposure data included 330,759 cases from the UK Biobank [26]. Outcome data included 1301 cases from the UK Biobank [27] and 424 cases alongside 217,185 controls from the FinnGen Biobank. (Supplementary Table S1) Ethical clearance was secured for each component study prior to conducting the Mendelian Randomization (MR) Analysis. The UK

Biobank, a forward-looking study, encompasses over half a million UK residents aged 40 to 69 at enrollment (2006-2010), and has amassed genome-wide genotyping and phenotypic data on a broad spectrum of traits.

2.3.2. Instrumental Variables

In this analysis, three sets of GWAS summary data for exposures were extracted (Supplementary Table S1). Single nucleotide polymorphisms (SNPs) within the designated loci, exhibiting significance levels of 5×10^{-8}, were selected as instrumental variables. A fundamental principle of Mendelian Randomization (MR) dictates the absence of linkage disequilibrium (LD) among the chosen instrumental variables to preclude the emergence of biased outcomes. Accordingly, to evaluate LD among SNPs, cluster processing was undertaken with a stringent threshold ($R^2 < 0.001$ and a cluster distance of 10,000 kb), ensuring the reliability and integrity of the instrumental variables employed in this study.

2.3.3. Hypotheses

This two-sample Mendelian Randomization (MR) study is predicated on three foundational assumptions to mitigate bias. Firstly, the genetic instruments employed exhibit significant associations with the exposure. Secondly, the instrumental variables are not correlated with confounders influencing both exposure and outcome, ensuring their independence. Lastly, the instrumental variables exert influence on the outcomes solely via the exposure pathway, indicating the absence of horizontal pleiotropy wherein the instrumental variables would have effects on the outcomes that are not mediated through exposure [28].

2.3.4. Data Analysis

A variety of analytical techniques were employed to explore potential causal inferences, encompassing inverse variance weighting (IVW) [29], MR-Egger, weighted mode [30], weighted median [31], and simple mode [32]. In scenarios devoid of horizontal pleiotropy, IVW findings were deemed robust [33]. Heterogeneity was assessed utilizing the 'mr_heterogeneity' function within the "TwoSampleMR" R package, applying Cochran's Q test for the IVW method and MR-Egger regression. Horizontal pleiotropy was examined through the 'mr_pleiotropy_test' function, also in the "TwoSampleMR" package, leveraging the MR-Egger approach. MR-Egger regression was then used to estimate the impact of pleiotropy, providing more reliable causal estimates after adjusting for pleiotropy under the presumption of no measurement errors and instrument strength being independent of direct effects [34]. Should MR-Egger indicate pleiotropy, MR-PRESSO [35] was employed to adjust for outliers. Leave-one-out analysis determined the causal influence of individual SNPs. The *F*-statistic was calculated to ascertain the presence of weak instrument bias within the chosen instrumental variables, with $F > 10$ suggesting the absence of such bias, thereby reinforcing the hypothesis of association. The formula for the *F*-statistic is $F = \frac{R^2 * (N-2)}{1-R^2}$, where 'N' represents the sample size of the exposure, and 'R^2' denotes the proportion of exposure variation explained by the instrumental variable [36].

2.4. Construction of Machine Learning Model

Initially, data from NHANES underwent a cleaning process to extract relevant feature indicators and predictors. This process involved segregating features into continuous and categorical variables, applying one-hot encoding to categorical variables, and scaling continuous variables for normalization [37].

Within the Python (Version 3.8) programming environment, the Scikit-learn library [38] was utilized to construct nine machine learning models: Support Vector Machine (SVM), Random Forest Classifier (RF), Logistic Regression Classifier (LR), Gradient Boosting Classifier (GB), K-Nearest Neighbors Classifier (KNN), Light Gradient Boosting Machine (LGBM), CatBoost (CATB), Decision Tree (DT), and AdaBoost (AD). Hyperparameter optimization for these models was conducted using

GridSearchCV, a method that combines grid search with cross-validation, to identify and store each model's optimal settings. This optimization process employed a ten-fold cross-validation approach to assess model performance, ultimately selecting the combination of parameters that exhibited the best performance.

Following the principle of random stratified sampling, the dataset was partitioned into a training set and a test set in an 8:2 ratio. Subsequently, nine machine learning models were trained using this data partitioning.

2.5. Evaluation of Machine Learning Models

Utilizing the optimized parameters, the models were trained, and their performance metrics were computed, including the generation and comparison of classification reports and confusion matrices [39]. Precision, recall, accuracy, F1 score, and the area under the receiver operating characteristic (ROC) curve (AUC) served as the evaluation criteria for the machine learning models [40]. Additional analyses, such as precision-recall, decision boundary analysis, and Bootstrap ROC analysis, were conducted to further assess model performance [39].

The Shapley Additive exPlanation (SHAP), a framework introduced by Lundberg and Lee [41] for elucidating machine learning predictions, offers a novel approach to interpreting various complex models, providing interpretable insights that has been substantiated in prior research [42]. In our study, SHAP values were employed to highlight critical features for the prediction of early-stage CKD, facilitating an examination of the pivotal factors that influence outcome predictions. This analysis delineated the impact of each significant feature on the predictive performance of the final machine learning model, enhancing our understanding of the model's decision-making process.

2.6. Creation of Interactive Web Interface

A web application was developed to provide online access to our novel prediction model, utilizing the "shiny" R package as the foundational framework. The application can be accessed at https://guanmiao.shinyapps.io/machinelearningckd88/. Furthermore, the input variables were quantified. (Supplementary Table S2)

3. Results

3.1. Baseline Characteristics

A cohort comprising 12,392 individuals was recruited, with an average age of 43.3 years; among these, 2,060 were diagnosed with CKD while 10,332 were not, as summarized in Table 1. Preliminary analysis of the baseline data revealed that the mean age of patients with CKD was significantly greater compared to those without CKD, with the majority of participants being non-Hispanic white. Regarding gender distribution, the non-CKD group exhibited a balanced ratio of male to female participants, whereas the CKD group had a predominance of female patients. Furthermore, notable disparities were observed between the CKD and non-CKD groups in terms of demographic characteristics, including economic status, race, citizenship, level of education, and household size, with all differences achieving statistical significance ($P < 0.05$).

Table 1. Demographic information analysis.

Variable	No (N = 10,332)	Yes (N = 2,060)	*P* value
age	41.50 (40.94,42.05)	52.09 (50.67,53.51)	< 0.0001
poverty	3.03 (2.94,3.13)	2.71 (2.59,2.84)	< 0.0001
sex			< 0.0001
Female	5300 (51.13)	1158 (59.72)	
Male	5032 (48.87)	902 (40.28)	
eth1			< 0.001
Mexican American	2067 (8.84)	342 (8.07)	

Non-Hispanic Black	2166 (9.97)	533 (14.23)	
Non-Hispanic White	4134 (68.92)	878 (66.30)	
Other Hispanic	855 (5.87)	138 (5.02)	
Other Race including multi-racial	1110 (6.41)	169 (6.38)	
Country_of_birth			0.01
Born Elsewhere	350 (4.82)	59 (4.68)	
Born in 50 U.S. states or Washington DC	3467 (39.08)	687 (37.44)	
Born in 50 U.S. States or Washington DC	4173 (44.61)	981 (48.81)	
Born in Mexico	611 (2.55)	81 (1.53)	
Born in Other Non-Spanish Speaking Country	67 (0.41)	24 (0.80)	
Born in Other Spanish Speaking Country	44 (0.15)	16 (0.32)	
Others	1620 (8.39)	212 (6.42)	
citizenship			< 0.0001
Citizen by birth or naturalization	8779 (90.36)	1868 (94.10)	
Not a citizen of the U.S.	1553 (9.64)	192 (5.90)	
household_size			< 0.0001
1	961 (10.05)	377 (16.54)	
2	2619 (31.08)	661 (37.90)	
3	1887 (19.01)	295 (14.61)	
4	2025 (20.19)	282 (13.67)	
5	1408 (11.15)	236 (10.28)	
6	691 (4.44)	108 (3.89)	
7 or more people in the Household	741 (4.08)	101 (3.11)	
9-11th grade (Includes 12th grade with no diploma)	505 (3.44)	114 (4.39)	
9-11th Grade (Includes 12th grade with no diploma)	539 (4.86)	169 (7.90)	
9th grade	66 (0.41)	6 (0.16)	
9th Grade	192 (0.73)	31 (1.13)	
College graduate or above	1313 (15.90)	149 (11.29)	
College Graduate or above	798 (12.75)	122 (8.21)	
GED or equivalent	2 (0.01)	0 (0.00)	
GED or Equivalent	16 (0.06)	2 (0.05)	
High School Grad/GED or Equivalent	817 (11.23)	207 (12.22)	
High school graduate	69 (0.46)	8 (0.25)	
High School Graduate	136 (0.60)	15 (0.54)	
High school graduate/GED or equivalent	971 (8.77)	193 (9.42)	
Less Than 5th Grade	0 (0.00)	1 (0.02)	
Less than 9th grade	377 (1.88)	80 (2.21)	
Less Than 9th Grade	443 (2.54)	170 (5.41)	
More than high school	153 (0.78)	23 (0.83)	
Never attended / kindergarten only	0 (0.00)	9 (0.40)	
Some college or AA degree	1379 (14.07)	232 (11.32)	
Some College or AA degree	980 (14.35)	196 (13.07)	

In the self-reported data (Table 2), 84.1% of participants indicated either no perceived difference in hearing between their ears or were unaware of any disparity. The analysis of otoscopic screening

data revealed a statistically significant variation in the incidence of excessive cerumen between the left and right ears among individuals with and without CKD, and the proportion of patients with cerumen excess in the CKD group was higher than that in the non-CKD group. Furthermore, significant discrepancies were noted in the outcomes of normal otoscopic examinations between the two groups (P<0.05), with a greater proportion of individuals in the non-CKD group exhibiting normal otoscopic findings compared to those in the CKD group.

Table 2. Analysis of otological status indicators.

Variable	Total	No	Yes	*P* value
self_reported_better_ear				0.03
no / don't know	10422 (84.1)	8766 (83.73)	1656 (80.76)	
yes, left ear	832 (6.71)	656 (7.10)	176 (8.82)	
yes, right ear	1138 (9.18)	910 (9.17)	228 (10.42)	
excessive_cerumen_left_ear				< 0.0001
no	10971 (88.53)	9227 (90.88)	1744 (86.56)	
yes	1421(11.47)	1105 (9.12)	316 (13.44)	
impacted_cerumen_left_ear				0.03
no	12089 (97.55)	10072 (97.56)	2017 (98.53)	
yes	303 (2.45)	260 (2.44)	43 (1.47)	
otoscopy_left_ear				0.002
no	1798 (14.51)	1423 (12.13)	375 (15.94)	
yes	10594 (85.49)	8909 (87.87)	1685 (84.06)	
collapsing_ear_canals_left_ear				0.11
no	12305 (99.3)	10271 (99.29)	2034 (98.84)	
yes	87 (0.7)	61(0.71)	26 (1.16)	
normal_otoscopy_right_ear				< 0.0001
no	1957 (15.79)	1550 (13.45)	407 (18.61)	
yes	10435 (84.21)	8782 (86.55)	1653 (81.39)	
excessive_cerumen_right_ear				< 0.0001
no	10862 (87.65)	9131 (89.81)	1731 (84.44)	
yes	1530 (12.35)	1201 (10.19)	329 (15.56)	
impacted_cerumen_right_ear				0.84
no	12045 (97.2)	10049 (97.36)	1996 (97.46)	
yes	347 (2.8)	283 (2.64)	64 (2.54)	
collapsing_ear_canals_right_ear				0.26
no	12313 (99.36)	10276 (99.32)	2037 (99.00)	
yes	79 (0.64)	56 (0.68)	23 (1.00)	yes

AUX_G Hearing data file is displayed in Table 3, (1) Analysis of bilateral tympanic chamber measurement data: there were significant differences in tympanic volume and compliance between CKD group and CKD group ($P < 0.05$), but there was no statistical significance in middle ear pressure and tympanic width between subjects with and without CKD ($P > 0.05$). (2) Analysis of hearing threshold data of pure tone air conduction measurement: The test thresholds of the left and right ears of the CKD group and the non-CKD group at all frequencies are significantly different, and the test results all show that the average hearing threshold of the CKD group is higher than that of the non-CKD group, that is, the average hearing loss of the CKD group is greater in different frequency groups.

Table 3. Analysis of otological status indicators.

Variable	No (N = 10,332)	Yes (N = 2,060)	*P* value
tympanic_right_middle_ear_pressure_dapa	-12.22 (-13.23,-11.22)	-11.58 (-13.62, -9.54)	0.57
tympanic_right_physical_volume_cc	1.44 (1.42,1.46)	1.39 (1.36,1.42)	0.003

tympanic_right_width	82.08 (80.97,83.19)	83.61 (81.89,85.34)	0.14
tympanic_right_compliance	0.79 (0.78,0.81)	0.72 (0.69,0.75)	< 0.0001
tympanic_left_middle_ear_pressure_dapa	-11.15 (-12.19,-10.10)	-9.86 (-11.50, -8.23)	0.23
tympanic_left_physical_volume_cc	1.40 (1.38,1.43)	1.34 (1.31,1.37)	< 0.001
tympanic_left_width	82.12 (81.00,83.24)	83.27 (81.27,85.26)	0.31
tympanic_left_compliance	0.79 (0.78,0.81)	0.72 (0.68,0.75)	< 0.001
right_threshold_1000hz_db	10.18 (9.79,10.57)	15.51 (14.57,16.46)	< 0.0001
right_threshold_500hz_db	11.30 (10.91,11.70)	16.15 (15.32,16.99)	< 0.0001
right_threshold_1000hz_2nd_read_db	9.61 (9.23,10.00)	15.17 (14.21,16.12)	< 0.0001
right_threshold_2000hz_db	10.86 (10.39,11.34)	19.14 (17.87,20.42)	< 0.0001
right_threshold_3000hz_db	14.30 (13.77,14.83)	24.15 (22.66,25.64)	< 0.0001
right_threshold_4000hz_db	17.75 (17.10,18.39)	29.07 (27.46,30.68)	< 0.0001
right_threshold_6000hz_db	23.87 (23.06,24.67)	36.07 (34.45,37.68)	< 0.0001
right_threshold_8000hz_db	24.74 (23.81,25.67)	39.87 (37.93,41.81)	< 0.0001
left_threshold_1000hz_db	9.90 (9.57,10.23)	15.20 (14.29,16.11)	< 0.0001
left_threshold_500hz_db	11.10 (10.80,11.41)	15.71 (14.96,16.47)	< 0.0001
left_threshold_1000hz_2nd_read_db	9.56 (9.22, 9.89)	14.97 (14.08,15.86)	< 0.0001
left_threshold_2000hz_db	11.35 (10.90,11.81)	19.53 (18.36,20.70)	< 0.0001

3.2. Logistic Regression Analysis of Hearing Indicators

Upon examining the baseline data, we incorporated all pertinent data into the logistic regression analysis to identify factors contributing to CKD and their 95% CIs (Table 4). The logistic regression model revealed:

(1) A negative correlation between the tympanic volume and compliance in both ears and CKD, indicating that as tympanic volume and compliance decrease, the odds of CKD increase (OR < 1, $P <$ 0.05).

(2) No significant association was observed between tympanic pressure and width with CKD [OR = 1.00, 95% CI = (1.00, 1.00), $P >$ 0.05].

(3) Hearing thresholds at various frequencies in both ears exhibited a positive correlation with CKD, implying that higher hearing thresholds are associated with an increased risk of CKD [OR = 1.02/1.03, $P <$ 0.0001].

Table 4. Analysis of otological status indicators.

Character	Est.	Std. Err	*t* value	Pr (>|t|)	OR	95% CI
tympanic_right_middle_ear_pressure（dapa）	0	0	0.55	0.58	1	1.00 (1.00,1.00)
tympanic_right_physical_volume（cc）	-0.25	0.09	-2.89	0.004	0.78	0.78 (0.65,0.92)
tympanic_right_width	0	0	1.5	0.14	1	1.00 (1.00,1.00)
tympanic_right_compliance	-0.34	0.09	-3.7	<0.001	0.71	0.71 (0.59,0.85)
tympanic_left_middle_ear_pressure（dapa）	0	0	1.16	0.25	1	1.00 (1.00,1.00)
tympanic_left_physical_volume_cc	-0.35	0.1	-3.61	<0.001	0.71	0.71 (0.59,0.86)
tympanic_left_width	0	0	1.02	0.31	1	1.00 (1.00,1.00)
tympanic_left_compliance	-0.34	0.11	-3.14	0.002	0.71	0.71 (0.57,0.88)
right_threshold_1000hz_db	0.03	0	12.5	<0.0001	1.03	1.03

						(1.03,1.04)
right_threshold_500hz_db	0.03	0	11.77	<0.0001	1.03	1.03
						(1.03,1.04)
right_threshold_1000hz_2nd_read_db	0.03	0	13.32	<0.0001	1.03	1.03
						(1.03,1.04)
right_threshold_2000hz_db	0.03	0	15.15	<0.0001	1.03	1.03
						(1.03,1.04)
right_threshold_3000hz_db	0.03	0	15.99	<0.0001	1.03	1.03
						(1.02,1.03)
right_threshold_4000hz_db	0.02	0	16.85	<0.0001	1.02	1.02
						(1.02,1.03)
right_threshold_6000hz_db	0.02	0	16.2	<0.0001	1.02	1.02
						(1.02,1.03)
right_threshold_8000hz_db	0.02	0	18.54	<0.0001	1.02	1.02
						(1.02,1.02)
left_threshold_1000hz_db	0.03	0	12.92	<0.0001	1.03	1.03
						(1.03,1.04)
left_threshold_500hz_db	0.03	0	12.25	<0.0001	1.03	1.03
						(1.03,1.04)
left_threshold_1000hz_2nd_read_db	0.03	0	13.37	<0.0001	1.03	1.03
						(1.03,1.04)
left_threshold_2000hz_db	0.03	0	14.97	<0.0001	1.03	1.03
						(1.03,1.04)

3.3. Correlation Analysis

Table 5 presents the ten most significant features as delineated by their correlation coefficients. The magnitude of the coefficient's absolute value is directly proportional to the feature's impact on the model's predictive outcome. Specifically, a positive coefficient signifies a direct correlation with an increase in the feature value leading to a higher likelihood of the positive class, exemplified by the diagnosis of chronic kidney disease, whereas a negative coefficient denotes an inverse relationship. The elucidated correlation coefficients suggest that variables including citizenship, hearing threshold, and age possess considerable potential to influence the progression of CKD.

**Table 5.** Correlation analysis of characteristic index.

Feature	Correlation
citizenship	0.5542568
right_threshold_1000hz_db	0.4365804
self_reported_better_ear	0.3382538
poverty	0.2730662
age	0.2408413
tympanic_left_compliance	0.2054256
tympanic_left_physical_volume_cc	0.2003082
left_threshold_1000hz_2nd_read_db	0.1580396
right_threshold_3000hz_db	0.1485964
impacted_cerumen_left_ear	0.1237963

3.4. Mendelian Randomization Analysis

Following the clustering of SNPs and data harmonization, a total of 29 SNPs were identified as being associated with all outcomes. Table S3 encapsulates the $R2$ and F-statistics pertaining to exposure, with all exposures demonstrating F-statistics greater than 10. This absence of weak

instrument bias underscores the robust validity of all SNPs under consideration. (Supplementary Tables S3)

In our investigation of the causal relationships between certain chronic kidney disease factors and hearing impairment, as illustrated by scatter plots and forest plots (**Figure 2**), the results indicate a significant positive correlation. The Inverse Variance Weighted (IVW) method reveals a causal relationship between MTAG and kidney injury molecule levels, demonstrating a significant positive correlation (**Figure 2, A**) ($\beta = 1.116$, OR $= 3.051$, 95%CI: (1.064, 8.748), $P = 0.0379$). This finding suggests that an increase in hearing impairment is associated with elevated levels of kidney injury markers. Despite an overarching positive trend across all five analytical methods (inverse variance weighting, MR Egger, simple mode, weighted median, and weighted mode) (**Figure 2, B**), the wide confidence intervals introduce some uncertainty regarding the specific magnitude of this effect. Nonetheless, they uniformly affirm a statistically significant connection between MTAG and the kidney injury molecule levels. Furthermore, MTAG exhibited a causally significant and positive association with other specified disorders of kidney and ureter (Figure 2E) ($\beta = 2.160$, OR $= 8.673$, 95% CI: (1.427, 52.728), $P = 0.0190$), intimating that worsening hearing impairment is closely and significantly linked to an augmented risk of these conditions. While the five methodologies similarly indicate an overall positive correlation (**Figure 2**, F), their extensive confidence intervals reveal some ambiguity regarding the impact's definitive extent. (Supplementary Table S4)
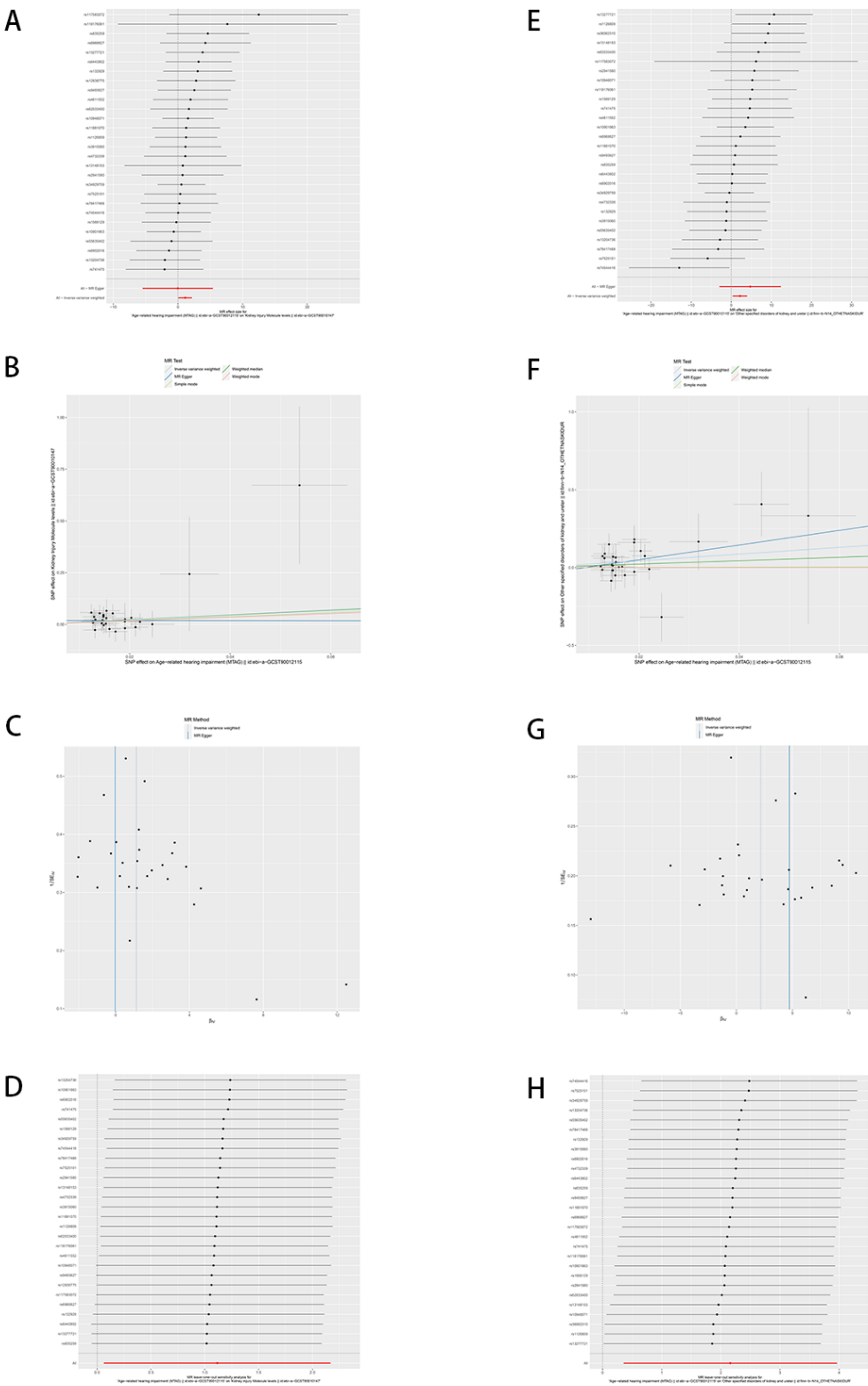
Figure 2. Result graph of Mendelian randomization. Mendelian randomization analysis of age-related hearing impairment (MTAR) and kidney injury Molecular levels included (A) forest plot, (B) scatter plot, (C) funnel plot, and (D) leave-one-out sensitivity analysis result plot. The Mendelian randomization analysis of MTAR and other specified disorders of kidney and ureter included (E) forest plot (F) scatter plot (G) funnel plot (H) leave-one-out sensitivity analysis results plot.

To assess the robustness of our study findings, further analyses were conducted on the included SNPs using MR-Egger and MR-PRESSO tests. These analyses did not detect any potential horizontal

pleiotropy (*P* > 0.05), (Supplementary Table S5) indicating an absence of pleiotropic effects among the SNPs. However, it is noteworthy that funnel plots revealed a slight asymmetry (**Figure 2, C and G**). Such asymmetry within the funnel plots suggests the presence of some degree of potential bias or heterogeneity among the study outcomes. While this does not directly undermine our findings, it prompts consideration of the possibility of unexplained minor biases or variations present in the data.

Adjusted Cochran Q statistics for the current analysis indicate significant heterogeneity among the SNP effects included in our study. (Supplementary Table S5) Furthermore, leave-one-out sensitivity analysis was employed to assess the impact of each individual SNP on the overall causal relationship. (**Figure 2, D and H**) As illustrated, the outcomes of this analysis demonstrate that systematically excluding each SNP and repeating the MR analysis does not result in any significant alterations to the observed causal relationship. This consistency across various iterations of the analysis reinforces the conclusion that the estimated effects are not driven by any single genetic variant. This finding substantiates the robustness of the causal inferences drawn from our data, suggesting that the genetic instruments constitute a well-distributed set that collectively contributes to the analysis.

3.5. Machine Learning model Construction

Nine machine learning models were developed, encompassing Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), Gradient Boosting (GB), K-Nearest Neighbors (KNN), Light Gradient Boosting Machine (LGBM), CatBoost (CATB), Decision Tree (DT), and AdaBoost (AD). Following the optimization of hyperparameters, the optimal settings for each model were determined. (Supplementary Table S6)

3.6. Machine Learning Model Comparison and Evaluation

We embarked on a comparison and evaluation of various machine learning models for the binary classification of CKD, utilizing two distinct performance assessment methodologies. Initially, the Receiver Operating Characteristic (ROC) curves during the training process were plotted (**Figure 3**). This figure displays the ROC curve for each model, with the Area Under the Curve (AUC) scores provided in the legend. The AUC score quantifies the two-dimensional area beneath the entirety of the ROC curve, offering a comprehensive measure of performance across all classification thresholds. The proximity of the ROC curve to the top left corner of the chart signifies greater test accuracy. Correspondingly, an AUC score nearing 1 indicates the model's proficiency in predicting positive cases while minimizing false positives. The figure illustrates our ROC curves closely aligned with the upper left corner, denoting high accuracy. CatBoost emerged with the highest AUC at 0.931, marking it as the most effective among the evaluated models. In contrast, the K-Nearest Neighbors (KNN) model had the lowest AUC at 0.860, reflecting its poor trade-off between true positive and false positive rates. Subsequently, recall and decision boundary graphs were delineated (Figure 3B), comprising separate charts for the CKD and non-CKD classes. Each line within these graphs represents the recall rate of distinct models across various decision boundary thresholds (T). For the CKD class (left graph), recall rates ascend with increasing decision boundaries, whereas for the non-CKD class (right graph), most models initially exhibit high recall rates, which then progressively diminish. This pattern indicates a trade-off in accurately classifying CKD versus non-CKD cases. The ideal threshold for balancing recall rates between these categories lies near the convergence or closest approach of the curves. The adjustment of recall rates is tailored to the specific requirements of the task at hand. CatBoost, Logistic Regression, and Light Gradient Boosting models notably achieve an optimal balance in overall performance.
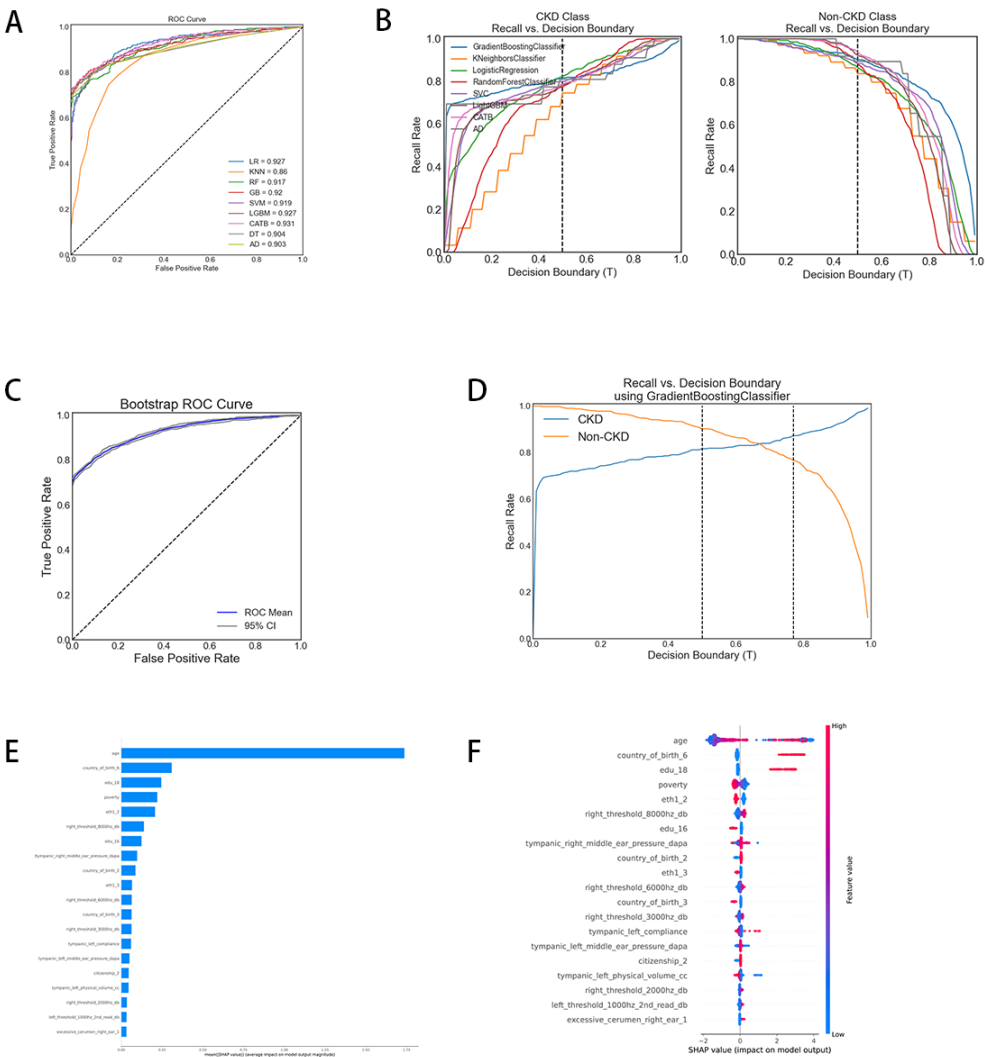
Figure 3. (A) ROC curves of nine models; (B) Recall and decision boundary diagram of nine models (C) ROC curve of CatBoost model; (D) Recall and Decision boundary graph of CatBoost model (E) Based on mean (|SHAP value |), feature ranking (Y-axis) indicates importance of prediction model. The SHAP value (X-axis) is a uniform index that responds to the influence of a feature in the model. ; (F) Ranking the stability and importance of interpretation of the top 20 features using an optimal model. In each feature important row, all patients' attributions to the outcome are plotted with different colored dots, where red dots represent high risk values and blue dots represent low risk values.

Subsequently, we compared various metrics across nine machine learning models (Table 6). Accuracy, defined as the ratio of correct predictions to total predictions, indicates the model's efficacy in correctly predicting labels. A higher accuracy denotes a greater number of correctly predicted labels. According to the results, the CatBoost model achieved the highest accuracy (0.862), signifying it had the highest percentage of correctly classified instances. Conversely, the KNN model exhibited the lowest accuracy, indicating the highest number of classification errors.

Regarding the AUC, the CatBoost model demonstrated superior performance, closely followed by Logistic Regression and Light Gradient Boosting Machine, suggesting these models excel in distinguishing between categories. When assessing accuracy, the CatBoost model also ranked highest, with Decision Tree and Gradient Boosting models trailing closely behind.

Further analysis of model performance, as per Table 6, reveals that among all models, CatBoost exhibits an exceptionally high recall rate for the non-CKD class and the best precision for the CKD class, with high F1 scores for both categories. CatBoost appears to maintain the optimal overall

balance between accuracy and recall across these two categories. The KNN model underperforms in predicting CKD, manifesting the lowest F1 score for this category. The Decision Tree model shows a very high recall for the non-CKD class, indicating a potential overprediction of this category at the expense of precision for the CKD class. When considering both macro and weighted averages, CATB also performs well in terms of accuracy, recall and F1 scores. It was positioned as a potentially superior model for this dataset based on these metrics. GB and DT also demonstrate strong performance, albeit slightly inferior to CATB.

Consequently, considering its superior overall performance, we have selected CatBoost as our preferred model, characterized by an AUC of 0.931 and an accuracy of 0.862. The stability and discriminative capacity of the model are further corroborated by the Bootstrap ROC curve (**Figure 3, C**), with the 95% confidence interval closely hugging the mean line, indicating consistent model performance across different samples. The recall versus decision boundary graph (**Figure 3, D**) illustrates that within the threshold range of 0.5 to 0.75, the model achieves an optimal balance between identifying CKD cases and maintaining a low false-positive rate. These findings underscore the CatBoost model's capability to accurately identify positive cases while minimizing misdiagnoses, making it particularly well-suited for scenarios requiring precise identification of CKD cases.

Subsequently, we employed SHAP values to elucidate how various variables influence CKD predictions within the model. Figure 3E illustrates the top 20 risk factors assessed by average absolute SHAP values. Figure 3F highlights the 20 most significant features within our model. It was observed that age stands out as the most critical feature, exerting a substantial average impact on the model's output magnitude. Additionally, several demographic factors, such as place of birth, education level, poverty status, and ethnicity, also significantly influence the model. Hearing-related variables, like "right_threshold_8000hz_db" and "tympanic_right_middle_ear_pressure_dapa," play an important role in the model's decision-making process.

Ultimately, we have established a website that utilizes the CatBoost algorithm to predict CKD from auditory and demographic data, available at (https://guanmiao.shinyapps.io/machinelearningckd88/). This platform features a user-friendly interface that permits the entry of clinical characteristics for new samples. This web application also offers predictions regarding the likelihood of CKD presence or absence, tailored to the specifics of the user-provided information.

4. Discussion

In this study, we explored interpretable ML approaches utilizing auditory and demographic indicators from the US NHANES data spanning from 2000 to 2020, aimed at identifying CKD. The study encompassed 12,392 participants, including 2,060 individuals diagnosed with CKD. By comparing the baseline characteristics of the CKD and non-CKD groups, we observed that individuals with CKD were, on average, older and that there was a higher proportion of females among these patients. Additionally, significant correlations were identified between CKD and various demographic factors such as economic status, race, citizenship, level of education, and household size, underscoring the importance of considering demographic characteristics in CKD research and management. Most participants perceived their hearing in both ears to be similar. However, otoscopic examinations revealed significant differences between the CKD and non-CKD groups in terms of cerumen excess and otoscopic results. Furthermore, a positive correlation was observed between hearing thresholds and CKD, indicating greater average hearing loss within the CKD group. These results suggest that auditory indicators may serve as critical health markers for individuals with CKD. Through Mendelian randomization analysis, we discovered causal relationships between certain factors of CKD and hearing impairment. Significantly, our research elucidated substantial causal associations between other specified disorders of kidney and ureter and age-related hearing impairment, as well as establishing the causality between kidney injury molecule levels and MTAG. These findings, based on extensive clinical data, not only reinforce the potential biological connection between CKD and hearing but also represent a pioneering investigation into the causal relationship between factors associated with CKD and hearing impairment. Additionally,

they provide a theoretical foundation for ML models that use auditory and demographic indicators for the early prediction of CKD.

Among nine ML models, the CATB model achieved the best performance, ultimately being selected to identify CKD. Tested across various metrics, the CATB model demonstrated an average AUC of 0.931, indicating efficient and stable classification capabilities, coupled with a high accuracy of 0.862, as well as superior F1-score, recall, and other metrics. Nevertheless, different models possess distinct advantages; some may excel in identifying CKD patients, while others may be better at confirming non-CKD cases. The choice of model might depend on clinical requirements. If minimizing false negatives is paramount, models with a high recall rate for the CKD category at the selected threshold would be preferable. Conversely, if reducing false positives is more critical, models with a high recall rate for the non-CKD category would be favored. The CATB model demonstrates robust and stable performance in classifying diseases such as CKD, striking an excellent balance between identifying actual cases and avoiding false positives. Yet, the precise balance between recall and precision can be adjusted according to the specific costs associated with false negatives and false positives in the application domain, allowing for fine-tuning of decision thresholds to achieve the balance required for practical use [43].

In Table 7, we catalog a range of machine learning models previously employed for predicting CKD. Comparative analysis reveals that our model exhibits a higher AUC, denoting a superior capability in distinguishing between CKD and non-CKD instances. This elevated discriminative proficiency not only underscores the model's robustness but also its reliability in generating predictions. Moreover, the balanced performance of our model across accuracy, recall, and F1 scores highlights its efficacy in managing the binary classification of CKD and non-CKD categories. Notably, a high recall rate for the positive category ensures that the model rarely overlooks actual cases of CKD, a critical requirement for medical diagnostics. The considerable size of our dataset offers multiple advantages. It mitigates the risk of overfitting, a common challenge in machine learning models, thereby enhancing the model's ability to generalize well to new data. Additionally, it addresses issues related to class imbalance, a frequently encountered concern in medical datasets where one category may disproportionately represent. The extensive dataset also diminishes the impact of noise or incorrect data, contributing to the robustness of our predictive model.

Table 7. Analysis of otological status indicators.

Authors	Dataset	Algorithms	Performance	Limitations
Stefan Ravizza et al [19]	IBM Explorys & Indiana Network for Patient Care	Feature selection strategy and logistic regression	The overall AUC value is 0.7937	Prediction of CKD as a diabetic complication
Jiamin Chen et al [20]	MIMIC-IV	LASSO multivariate logistic regression analysis	The AUC value is 0.771	Only indicators for 14 attributes included, with potential selection bias
Dai Su et al [21]	HABCS	LR Lasso RF GBDT SVM DNN	The maximum AUC is 0.75	Limitations include a small dataset size and potential biases due to reliance on self-reported medical history.
Jing Xiao et al [22]	Shanghai Huadong Hospital	LR Ridge Lasso SVM RF KNN NN XGBoost	The highest AUC was 0.873	Only 5 population indicators, which predict 24-hour proteinuria.
Zheyi Dong et al [23]	People's Liberation Army (PLA) General Hospital	LightGBM XGBoost AdaBoost	The highest AUC is 0.815	Single-institution setting, small sample size, and potential

NN DT SYM LR	bias from missing data in EMR，not using suitable feature selection techniques for the dataset

CKD is defined as persistent abnormalities in kidney structure or function for more than three months [44]. Diagnosis primarily relies on routine screening or incidental findings through serum chemistry features and urinalysis, followed by confirmatory imaging studies (e.g., ultrasound, CT, MRI), renal biopsy, and histopathological examination [5]. Traditional machine learning (ML) models for early diagnosis of CKD typically utilize predictors derived from blood and urine tests along with demographic indicators. Our methodology introduces a pioneering approach by integrating auditory examination indicators with demographic predictors for CKD forecasting. These predictive markers are more readily accessible than traditional blood and urine analyses, requiring merely an ear examination and supplementary information gathered through a brief survey. Consequently, this predictive model can be broadly applied across various levels of medical institutions, from primary care to specialized clinics. Our research delves deeper into the association and causality between auditory indicators and CKD, pioneering a CKD prediction model based on auditory metrics. We assessed the significance of auditory indicators within the model using SHAP values, revealing their notable importance. This innovative approach to CKD detection offers enhanced convenience, cost-effectiveness, and accuracy, presenting a valuable instrument poised to advance early and broader CKD screening, particularly in environments where conventional methods are less accessible.

The model, leveraging data from the NHANES, principally mirrors the demographic characteristics of the U.S. population, including a higher proportion of African American and Hispanic individuals [24]. However, it does not comprehensively encompass global diversity, which may limit its applicability worldwide. The reliance on auditory examination metrics may overlook other critical medical indicators essential for the diagnosis of CKD. Clinical validation emerges as a pivotal next step, necessitating the testing of the model within clinical settings to ensure the reliability and adaptability of its predictions.

In summary, we have employed innovative otologic indicators in conjunction with demographic indicators as predictors, subsequently transforming them into a promising tool for the expedited screening of CKD. It is meritorious for future studies to test the model incorporating ear examinations in actual clinical settings, thereby enhancing its potential for wider application and adoption. More importantly, further investigations into whether interventions targeting these factors could reduce the incidence of CKD are warranted, thereby facilitating precise early-stage prevention and intervention.

5. Conclusions

Our investigation leverages NHANES data spanning from 2000 to 2020, with a focus on otologic indicators and demographic factors to identify CKD. The study elucidates the correlation and causation between CKD and auditory impairments, serving as the foundation for the development of a robust machine learning model. Notably, the CatBoost algorithm emerged as the most efficacious, exhibiting an average AUC of 0.931. Compared to traditional blood and urine tests, this novel approach herald new avenues for CKD screening, minimizing the risk of misdiagnoses or overlooked cases, and demonstrating potential for broader application across various medical settings. However, further clinical validation is required to ensure the predictive relevance of the model and its adaptability in real-world healthcare environments. This study synergizes traditional medical theories with contemporary technology, offering a simplified and more effective tool for the early detection of CKD.

6. Patents

There are no patents resulting from the work reported in this manuscript.

18

References

1. Chen, T.K.; Hoenig, M.P.; Nitsch, D.; Grams, M.E. Advances in the Management of Chronic Kidney Disease. *BMJ* **2023**, e074216.
2. Gansevoort, R.T.; Correa-Rotter, R.; Hemmelgarn, B.R.; Jafar, T.H.; Heerspink, H.J.L.; Mann, J.F.; Matsushita, K.; Wen, C.P. Chronic Kidney Disease and Cardiovascular Risk: Epidemiology, Mechanisms, and Prevention. *The Lancet* **2013**, *382*, 339–352.
3. Gargiulo, R.; Suhail, F.; Lerma, E.V. Hypertension and Chronic Kidney Disease. *Disease-a-Month* **2015**, *61*, 387–395.
4. Ling, J.; Ng, J.K.C.; Chan, J.C.N.; Chow, E. Use of Continuous Glucose Monitoring in the Assessment and Management of Patients With Diabetes and Chronic Kidney Disease. *Front. Endocrinol.* **2022**, *13*.
5. Kalantar-Zadeh, K.; Jafar, T.H.; Nitsch, D.; Neuen, B.L.; Perkovic, V. Chronic Kidney Disease. *Lancet* **2021**, *398*, 786–802.
6. Li, J.; Xu, X.; Luo, J.; Chen, W.; Yang, M.; Wang, L.; Zhu, N.; Yuan, W.; Gu, L. Choosing an Appropriate Glomerular Filtration Rate Estimating Equation: Role of Body Mass Index. *BMC Nephrol* **2021**, *22*, 197.
7. Liu, H.; Zhao, H.; Zheng, D.; He, W.; Liu, Y.; Jin, J.; He, Q.; Lin, B. Misdiagnosis of Chronic Kidney Disease and Parathyroid Hormone Testing during the Past 16 Years. *Sci Rep* **2023**, *13*, 15838.
8. Boateng, J.O.; Boafo, N.; Osafo, C.; Anim-Sampong, S. Hearing Impairment among Chronic Kidney Disease Patients on Haemodialysis at a Tertiary Hospital in Ghana. *Ghana Med J* **2019**, *53*, 197–203.
9. Kim, J.H.; Lee, D.-H.; Lee, B.; Lim, S.H.; Ahn, Y.H.; Kang, H.G.; Ha, I.-S.; Cheong, H.I. Renal Syndromic Hearing Loss Is Common in Childhood-Onset Chronic Kidney Disease. *J Korean Med Sci* **2020**, *35*, e364.
10. Shim, Y.J.; Choi, H.G.; Wee, J.H. Association between Chronic Kidney Disease and Sudden Sensorineural Hearing Loss: A Longitudinal Follow-Up Studies Using ICD-10 Codes in a National Health Screening Cohort. *J Clin Med* **2023**, *12*, 2861.
11. Wu, K.-L.; Shih, C.-P.; Chan, J.-S.; Chung, C.-H.; Lin, H.-C.; Tsao, C.-H.; Lin, F.-H.; Chien, W.-C.; Hsiao, P.-J. Investigation of the Relationship between Sensorineural Hearing Loss and Associated Comorbidities in Patients with Chronic Kidney Disease: A Nationwide, Population-Based Cohort Study. *PLoS One* **2020**, *15*, e0238913.
12. Sarin, V.; Sharma, A.; Chopra, I. High Frequency Hearing Loss in Chronic Renal Disease: A Cross-Sectional Study. *Indian J Otolaryngol Head Neck Surg* **2022**, *74*, 4046–4052.
13. Agrawal, M.; Singh, C.V. Sensorineural Hearing Loss in Patients With Chronic Kidney Disease: A Comprehensive Review. *Cureus 15*, e48244.

14. Rysz, J.; Gluba-Brzózka, A.; Franczyk, B.; Jabłonowski, Z.; Ciałkowska-Rysz, A. Novel Biomarkers in the Diagnosis of Chronic Kidney Disease and the Prediction of Its Outcome. *IJMS* **2017**, *18*, 1702.

15. ICD-10-CM Diagnosis Code N28.89 - Other Specified Disorders of Kidney and Ureter Available online: https://icdlist.com/icd-10/N28.89#google_vignette (accessed on 8 April 2024).

16. Titova, O.E.; Michaëlsson, K.; Larsson, S.C. Sleep Duration and Stroke: Prospective Cohort Study and Mendelian Randomization Analysis. *Stroke* **2020**, *51*, 3279–3285.

17. Davies, N.M.; Holmes, M.V.; Davey Smith, G. Reading Mendelian Randomisation Studies: A Guide, Glossary, and Checklist for Clinicians. *BMJ* **2018**, *362*, k601.

18. Skrivankova, V.W.; Richmond, R.C.; Woolf, B.A.R.; Yarmolinsky, J.; Davies, N.M.; Swanson, S.A.; VanderWeele, T.J.; Higgins, J.P.T.; Timpson, N.J.; Dimou, N.; et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *JAMA* **2021**, *326*, 1614–1621.

19. Ravizza, S.; Huschto, T.; Adamov, A.; Böhm, L.; Büsser, A.; Flöther, F.F.; Hinzmann, R.; König, H.; McAhren, S.M.; Robertson, D.H.; et al. Predicting the Early Risk of Chronic Kidney Disease in Patients with Diabetes Using Real-World Data. *Nat Med* **2019**, *25*, 57–59.

20. Chen, J.; Li, Y.; Liu, P.; Wu, H.; Su, G. A Nomogram to Predict the In-hospital Mortality of Patients with Congestive Heart Failure and Chronic Kidney Disease. *ESC Heart Failure* **2022**, *9*, 3167–3176.

21. Su, D.; Zhang, X.; He, K.; Chen, Y.; Wu, N. Individualized Prediction of Chronic Kidney Disease for the Elderly in Longevity Areas in China: Machine Learning Approaches. *Front. Public Health* **2022**, *10*, 998549.

22. Xiao, J.; Ding, R.; Xu, X.; Guan, H.; Feng, X.; Sun, T.; Zhu, S.; Ye, Z. Comparison and Development of Machine Learning Tools in the Prediction of Chronic Kidney Disease Progression. *J Transl Med* **2019**, *17*, 119.

23. Dong, Z.; Wang, Q.; Ke, Y.; Zhang, W.; Hong, Q.; Liu, C.; Liu, X.; Yang, J.; Xi, Y.; Shi, J.; et al. Prediction of 3-Year Risk of Diabetic Kidney Disease Using Machine Learning Based on Electronic Medical Records. *J Transl Med* **2022**, *20*, 143.

24. National Center for Health Statistics. *NHANES - About the National Health and Nutrition Examination Survey. [Accessed 2023 Dec 8].*;

25. The MRC IEU OpenGWAS Data Infrastructure. *bioRxiv* https://www.biorxiv.org/content/10.1101/2020.08.10.244293v1 (accessed on 29 March 2024).

26. Kalra, G.; Milon, B.; Casella, A.M.; Herb, B.R.; Humphries, E.; Song, Y.; Rose, K.P.; Hertzano, R.; Ament, S.A. Biological Insights from Multi-Omic Analysis of 31 Genomic Risk Loci for Adult Hearing Difficulty. *PLoS Genet* **2020**, *16*, e1009025.

27. Gilly, A.; Park, Y.-C.; Png, G.; Barysenka, A.; Fischer, I.; Bjørnland, T.; Southam, L.; Suveges, D.; Neumeyer, S.; Rayner, N.W.; et al. Whole-Genome Sequencing Analysis of the Cardiometabolic Proteome. *Nat Commun* **2020**, *11*, 6336.

28. Coneys, R.; Storm, C.S.; Kia, D.A.; Almramhi, M.; Wood, N.W. Mendelian Randomisation Finds No Causal Association between Urate and Parkinson's Disease Progression. *Mov Disord* **2021**, *36*, 2182–2187.

29. Burgess, S.; Butterworth, A.; Thompson, S.G. Mendelian Randomization Analysis with Multiple Genetic Variants Using Summarized Data. *Genet Epidemiol* **2013**, *37*, 658–665.

30. Bowden, J.; Davey Smith, G.; Burgess, S. Mendelian Randomization with Invalid Instruments: Effect Estimation and Bias Detection through Egger Regression. *Int J Epidemiol* **2015**, *44*, 512–525.

31. Bowden, J.; Davey Smith, G.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* **2016**, *40*, 304–314.

32. Lawlor, D.A.; Harbord, R.M.; Sterne, J.A.C.; Timpson, N.; Davey Smith, G. Mendelian Randomization: Using Genes as Instruments for Making Causal Inferences in Epidemiology. *Stat Med* **2008**, *27*, 1133–1163.

33. Yavorska, O.O.; Burgess, S. MendelianRandomization: An R Package for Performing Mendelian Randomization Analyses Using Summarized Data. *Int J Epidemiol* **2017**, *46*, 1734–1739.

34. Burgess, S.; Thompson, S.G. Interpreting Findings from Mendelian Randomization Using the MR-Egger Method. *Eur J Epidemiol* **2017**, *32*, 377–389.

35. Verbanck, M.; Chen, C.-Y.; Neale, B.; Do, R. Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred from Mendelian Randomization between Complex Traits and Diseases. *Nat Genet* **2018**, *50*, 693–698.

36. Papadimitriou, N.; Dimou, N.; Tsilidis, K.K.; Banbury, B.; Martin, R.M.; Lewis, S.J.; Kazmi, N.; Robinson, T.M.; Albanes, D.; Aleksandrova, K.; et al. Physical Activity and Risks of Breast and Colorectal Cancer: A Mendelian Randomisation Analysis. *Nat Commun* **2020**, *11*, 597.

37. Rodríguez, P.; Bautista, M.A.; Gonzàlez, J.; Escalera, S. Beyond One-Hot Encoding: Lower Dimensional Target Embedding. *Image and Vision Computing* **2018**, *75*, 21–31.

38. Scikit-Learn: Machine Learning in Python | The Journal of Machine Learning Research Available online: https://dl.acm.org/doi/10.5555/1953048.2078195 (accessed on 29 March 2024).

39. Fawcett, T. An Introduction to ROC Analysis. *Pattern Recognition Letters* **2006**, *27*, 861–874.

40.   Powers, D.M.W. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation 2020.
41.   A Unified Approach to Interpreting Model Predictions | Proceedings of the 31st International Conference on Neural Information Processing Systems Available online: https://dl.acm.org/doi/ 10.5555 /3295222.3295230 (accessed on 30 March 2024).
42.   Lundberg, S.M.; Erion, G.G.; Lee, S.-I. Consistent Individualized Feature Attribution for Tree Ensembles 2019.
43.   Davis, J.; Goadrich, M. The Relationship between Precision-Recall and ROC Curves. In Proceedings of the Proceedings of the 23rd international conference on Machine learning; Association for Computing Machinery: New York, NY, USA, June 25 2006; pp. 233–240.
44.   Levin, A.; Stevens, P.E.; Bilous, R.W.; Coresh, J.; Francisco, A.L.M.D.; Jong, P.E.D.; Griffith, K.E.; Hemmelgarn, B.R.; Iseki, K.; Lamb, E.J.; et al. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney International Supplements* **2013**, *3*, 1–150.