Annals of Epidemiology 113 (2026) 1–12

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.sciencedirect.com/journal/annals-of-epidemiology

Unravelling impact of comorbidities on mortality risks in CKD patients during the COVID-19 pandemic: An explainable AI-driven study

Zeinab Abdollahi [a], Lin Huo [b,*], Donald Fraser [c], Shang-Ming Zhou [d,*]

[a] School of Engineering, Computing, and Mathematics, University of Plymouth, Plymouth PL4 8AA, United Kingdom
[b] China-ASEAN School of Economics, Guangxi University, Nanning 530000, China
[c] Wales Kidney Research Unit, School of Medicine, Cardiff University, Cardiff CF14 4XN, United Kingdom
[d] Centre for Health Technology, Faculty of Health, University of Plymouth, Plymouth PL4 8AA, United Kingdom

ARTICLE INFO

Keywords:
COVID-19
Chronic kidney disease
Comorbidities
Mortality
Machine learning
Interpretability
Risk factors

ABSTRACT

Objectives: The chronic kidney disease (CKD) patients were at high risk for severe clinical complications during the COVID-19 pandemic. Our objectives were to evaluate comorbidity prevalence; predict mortality risks for CKD patients during the pandemic; assess how various health factors interact to influence mortality; and provide insights for targeted prevention strategies.

Method: We analysed data from 186,396 CKD patients in Mexico during the entire pandemic (Jan 2020- May 2023). Explainable artificial intelligence (XAI) methods with extreme gradient boosting (XGBoost) models and Shapley Additive Explanations (SHAP) were developed to predict mortality for CKD patients with model interpretations. Different metrics were used to comprehensively evaluate model's generalisation performances.

Results: The most prevalent comorbidities were hypertension (64.39 %), diabetes (49.79 %), and obesity (16.46 %). Male patients and older individuals showed higher risk for adverse outcomes. The overall mortality rate was 19.33 %, with significantly higher mortality in COVID-19 positive patients (33.9 %) compared to COVID-19 negative patients (10.1 %). Comorbidities with the most significant impact on the mortality included diabetes, hypertension, and obesity, which were more frequent in the COVID-19 positive group and associated with higher rates of intubation, and ICU admission. Pneumonia was identified as a major predictor of negative outcomes in CKD patients with COVID-19. CVD was more common in the COVID-19 negative group. Our machine learning models achieved performances of AUC= 0.76 and F_1-score= 0.75 for predicting mortality during the pandemic.

Conclusion: Targeted management of comorbid conditions, especially respiratory infections, is crucial in CKD patients during pandemics.

Introduction

Chronic kidney disease (CKD) is a multifaceted disorder occurring in approximately 13 % of the population worldwide and representing a significant health and financial burden, particularly among those with comorbidities [1,2]. The prevalence of CKD is rapidly increasing in Western countries, including the UK. Projections from 2012 estimate that by 2036, more than 4 million people in England will have CKD in stages 3–5 [1]. This trend is expected to worsen due to demographic shifts, the obesity epidemic, and the effects of climate change, all of which will likely contribute to increased kidney disease prevalence, with significant implications for global survival rates [3].

Several comorbidities, including hypertension, diabetes, cardiovascular disease (CVD), obesity, asthma, and chronic obstructive pulmonary disease (COPD), are commonly observed in patients with CKD. These conditions are known to exacerbate the severity of illnesses, potentially leading to increased mortality rate [4]. Unlike CVD, stroke and respiratory disease, CKD mortality continues to increase. Currently, CKD is the third fastest-growing cause of death globally and the only non-communicable disease with a persistent rise in age-adjusted

List of abbreviations: AUC, Area under the receiver operating characteristic curve; CVD, Cardiovascular disease; CKD, Chronic kidney disease; COPD, Chronic obstructive pulmonary disease; DKD, Diabetic kidney disease; IQR, Interquartile range; NPV, Negative predictive value; PPV, Positive predictive value; XAI, Explainable artificial intelligence; XGBoost, Extreme gradient boosting.

* Corresponding authors.
E-mail addresses: lhuo@gxu.edu.cn (L. Huo), shangming.zhou@plymouth.ac.uk, smzhou@ieee.org (S.-M. Zhou).

https://doi.org/10.1016/j.annepidem.2025.11.005
Received 21 November 2025; Accepted 23 November 2025
Available online 25 November 2025

mortality [5]. By 2040, CKD is expected to rank as the fifth leading cause of global reduced lifespan [6].

The global outbreak of COVID-19 has added an additional layer of complexity to the management of patients with renal disease, as evidence suggests that these individuals may be at higher risk for severe clinical complications when infected with the virus [7]. Patients on dialysis and kidney transplant recipients were especially vulnerable, often showing unusual symptoms and facing worse outcomes. CKD has also been identified as a major risk factor for COVID-19 mortality, with a direct link between the severity of kidney dysfunction and death rates [7]. However, there is limited research that specifically addresses how COVID-19 complications differ between renal patients with and without various comorbidities.

On the other hand, despite the growing prevalence and severity of CKD, public awareness remains low, highlighting the need for improved education and the implementation of sustainable strategies for risk stratification as critical public health priorities [3]. As the world continues to grapple with the aftermath of the COVID-19 pandemic, this study aims to fill these knowledge gaps by systematically analyzing the prevalence of comorbidities in CKD patients, both with and without COVID-19; developing and employing explainable AI (XAI) approach to assess how intricate interplay of various health factors affect mortality risks in CKD populations during the pandemic; and providing actionable insights into targeted prevention and treatment strategies for CKD patients in future pandemics or seasonable infectious diseases.

Materials and methods

This methodology combined traditional statistical approaches with advanced machine learning techniques and XAI, offering a robust framework for analyzing the complex relationships between chronic comorbidities, COVID-19, and mortality in renal patients. The general workflow of this study is summarized in Fig. 1.

Study population and data source

This retrospective study extracted anonymized health information for 186,396 patients with CKD from the publicly available national healthcare data General Directorate of Epidemiology in Mexico [8]. This dataset covered the entire COVID-19 pandemic period from 30 January 2020–5 May 2023 as declared by the WHO [9,10]. The data encompassed epidemiological information on suspected viral respiratory disease cases identified across the Health Sector, including both outpatient and inpatient cases. It covered patient demographics, epidemiology, chronic comorbidities, and clinical outcomes.

Key variables and outcomes

This study examined several key variables: 1) Demographic factors: age, gender; 2) Behavioural factors: smoking habits; 3) Clinical presentations: occurrence of pneumonia, necessity for endotracheal intubation; 4) Epidemiological factors: contact with other COVID-19 cases; 5) COVID-19 status: COVID-19 infection status. The healthcare authority defined a suspected case of COVID-19 as someone who had experienced at least two of these three symptoms within the last week: cough, fever, or headache, along with at least one of these additional symptoms: trouble breathing, joint pain, muscle aches, sore throat, runny nose, pink eye, or chest pain. Doctors then used a specific laboratory test called RT-PCR to confirm or rule out COVID-19; 6) Chronic comorbidities: COPD, bronchial asthma, hypertension, heart disease, diabetes, obesity, and other illnesses. Before COVID-19 testing, patients were asked about underlying health conditions that could increase the risk of serious complications if the patient tested positive for COVID-19.

Age was shown as medians with the 25th–75th interquartile range (IQR), as it was not normally distributed according to the Kolmogorov–Smirnov test and Q-Q plot. For further analysis, age was categorized into six distinct ranges: under 20, 21–30, 31–40, 41–50, 51–60, and over 60. Differences between groups were tested using the chi-square test for categorical variables and the *t*-test or Mann–Whitney *U* test for continuous variables, based on the normality of the data. Missing values were excluded from the statistical analysis.

The adverse clinical outcome of interest in this study was mortality. The factors for mortality prediction included *Age, Sex, Intensive Care Unit ("ICU"), Intubation, Pneumonia, Diabetes, Chronic Obstructive Pulmonary Disease ("COPD"), Asthma, Immunosuppression("Immunosup"), Hypertension, Other Disease ("OtherDisease"), Cardiovascular Disease ("CVD"), Obesity, Smoker, Contacting Other Covid-Infected Individuals ("ContactOtherCovid"), COVID-19 Infection ("Covid19_status")*. We then used these datasets to develop machine learning model aimed at predicting the adverse risk of mortality.

Machine learning model

We employed XGBoost algorithm, a decision tree-based ensemble model, for our predictive analysis. XGBoost was chosen for its several advantages: built-in regularization techniques to prevent overfitting;

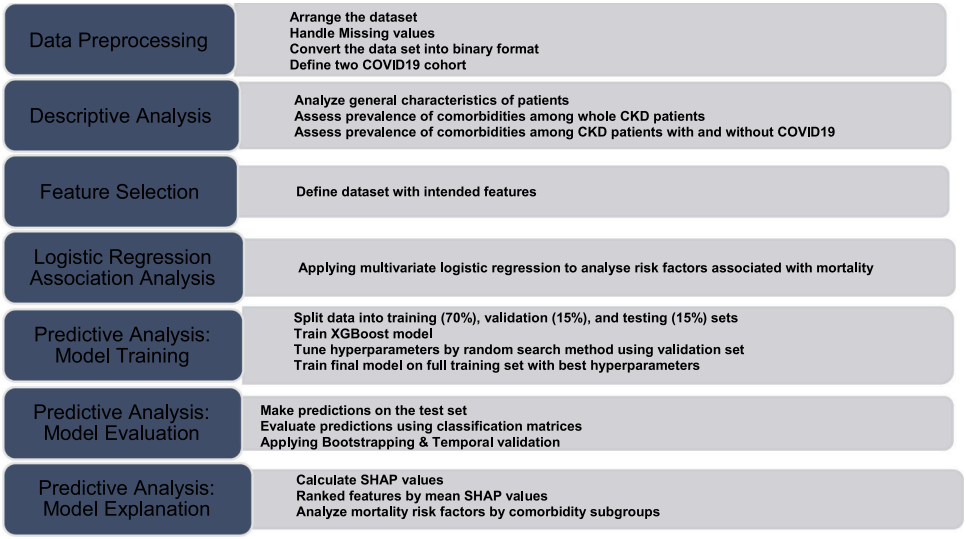

Fig. 1. General workflow of this study.

internal cross-validation function, simplifying model evaluation by removing the need for external packages [11]; flexibility in addressing various problem types (including regression, classification, and ranking) [12]; efficient data matrix and model storage capabilities, saving time on large datasets by avoiding redundant computations; and advanced tree pruning strategy, growing trees to a maximum depth and then pruning backwards until improvements in the loss function fall below a set threshold, leading to more optimal models [11].

We created XGBoost model for mortality prediction. The dataset was split into three parts: 70 % for training, 15 % for validation, and 15 % for testing. The training data was used to build the model with a 5-fold cross validation to tune the hyperparameters. Validation data was used to detect overfitting by providing an unbiased evaluation of the model's performance during training, and the test data was used to evaluate how well the model could generalize to new, unseen data.

In model development, the '*scale_pos_weight*' parameter was utilized to handle data imbalance. Random search method was employed for hyperparameter optimization due to its efficiency in exploring large hyperparameter spaces. Table M1 (in the Supplementary Materials) illustrates the XGBoost hyperparameters tuned in this study.

Explainable artificial intelligence (XAI)

XGBoost model captures complex connections and interactions in the data, thereby bolstering its efficacy in predictive tasks. However, this complexity can obscure the specific contributions of individual features to predictions. To enhance model interpretability, we employed explainable SHAP (SHapley Additive exPlanations) values, a post-hoc, model-agnostic method of XAI [13]. SHAP values quantify feature contributions to predictions both globally and locally, helping to explain complex models like XGBoost.

SHAP values can offer a global interpretability about the overall behaviour of the model, illustrating how the model makes decisions across the entire dataset and the significance of different features. At the same time, SHAP values can provide a local interpretability on individual predictions, explaining why a specific outcome was reached by examining the contribution of each feature to that prediction [14]. To interpret the contribution of individual comorbidities and their joint effects on predicted mortality, we applied SHAP to the final predictive XGBoost mortality model. SHAP values quantify the marginal contribution of each comorbidity to mortality potential for each subgroup, expressed in log-odds units of mortality prediction. These values decompose each feature's total SHAP contribution into an additive main effect and a pairwise interaction component, allowing assessment of whether the presence of one condition modifies the impact of another within the model.

For our binary features, interaction values were visualized as SHAP interaction dependence plots, which display four distinct regions representing all combinations of feature states (0/0, 1/0, 0/1, 1/1). The color scale in each plot reflects the interaction SHAP magnitude, where higher positive values indicate a synergistic increase in model-predicted mortality risk and negative values indicate attenuation of risk.

In addition, local SHAP waterfall plots and interaction dependency plots were generated for each comorbidity subgroup to visualize and interpret the feature-level contributions to mortality prediction at subgroup levels. These visualizations allowed detailed examination of how specific comorbidities influenced model output both independently and through their interactions with other risk factors.

Model evaluation

The model's performance was assessed using several metrics, including accuracy, the area under the receiver operating characteristic curve (AUC), F1-score, sensitivity (also known as recall), specificity, positive predictive value (PPV, also known as precision), and negative predictive value (NPV).

An AUC of 1.0 indicates the model performs perfectly, while an AUC of 0.5 means the model performs as well as random chance. Differently, F1-score is a measure of the harmonic mean of precision and recall.

$$F_1 - score = 2\frac{precision \bullet recall}{precision + recall}$$

F1-score can range from 0 to 1, with 1 representing a model that correctly classifies every observation into the appropriate class, and 0 representing a model that cannot classify any observations into the correct class.

Using multiple metrics, we aimed to provide a comprehensive evaluation of the models' predictive capabilities, balancing between different aspects of performance such as correctly identifying positive cases, minimizing false positives, and overall discriminative ability.

In order to mitigate the risk of overfitting and enhance the reliability of model evaluations, we performed bootstrapping with 1000 resamples to generate performance metrics with 95 % confidence intervals, thereby showing the precision of our model

Statistical analysis

The frequency of comorbidities was evaluated within the entire CKD cohort, assessing the prevalence and distribution of these conditions within the cohort. The prevalence of chronic conditions also was evaluated in renal patients both with and without COVID-19.

Moreover, we calculated the monthly mortality rate among COVID-19 positive (MDR_+) and negative (MDR_-) groups as follows

$$MDR_+ = \frac{NDM_+}{TNM_+}, \quad MDR_- = \frac{NDM_-}{TNM_-}$$

where NDM_+ is number of deaths among COVID-19 positive patients in the month, TNM_+ is total number of COVID-19 positive patients in the month, while NDM_- is number of deaths among COVID-19 negative patients in the month, TNM_- is total number of COVID-19 negative patients in the month.

Furthermore, we performed logistic regression to evaluate the strength of association of each factor with mortality as a comparison with XGBoost. Associations are presented as odds ratios (ORs) with 95 % confidence intervals. A p-value < 0.05 was considered statistically significant.

Software

All statistical analyses and machine learning model creation were performed using R software version 4.3.1. R packages for this study included 'caret', 'xgboost', 'SHAPforxgboost', 'shapviz'.

Results

Demographic characteristics

The 186,396 patients with CKD in this study included 87,837 females and 98,559 males. The mean age of the cohort was 53.85 years. The age distribution across different age groups can be found in Figure E4 in Supplementary Materials, which highlights a significant proportion of older patients, with 42 % of the cohort aged over 60 years.

Prevalence of comorbidities

As shown in Table 1, the most prevalent comorbidities were: hypertension (64.39 %), diabetes (49.79 %), obesity (16.46 %), cardiovascular disease (11.65 %), COPD (5.9 %). These findings underscore the high burden of metabolic and cardiovascular comorbidities in renal patients.

Table 1
Prevalence of different conditions in the entire renal patient population.

Condition	Number of Patients with Condition	Number of Patients without Condition	Missing Data	Prevalence (%)
CVD	21705	164522	169	11.65
Diabetes	92712	93469	215	49.79
Hypertension	119944	66309	143	64.39
Obesity	30656	155569	171	16.46
Asthma	5346	180902	148	2.87
COPD	10992	175202	202	5.90
Intubation	7551	83657	95188	8.27
ICU Admission	4254	86948	95194	4.66
Hospital Admission	91694	94702	-	49.19
Smoker	17740	168307	349	9.53
Pneumonia	47564	137375	1457	25.71
Immunosuppression	14006	172209	181	7.52

COVID-19 status analysis

After excluding patients with unspecified COVID-19 status the remaining patients were categorised into two groups: COVID-19 positive: 71,204 patients (39.9 %); COVID-19 negative: 107,258 patients (60.1 %) (Figure E5 in Supplementary Materials). Gender distribution showed that in females, 32,319 patients were COVID-19 positive, and 51,869 COVID-19 negative; in males, 38,885 patients were COVID-19 positive, and 55,389 COVID-19 negative. This distribution suggested a higher COVID-19 positivity rate among males (41.2 %) compared to females (38.4 %) (Figure E6 in Supplementary Materials). Moreover, the age distribution showed the median 57 years for the COVID-19 positive group and 55 years for the COVID-19 negative group (Figure E7 in Supplementary Materials).

Table E1 (in Supplementary Materials) showed the frequency of comorbidities and other variables in people who tested positive and negative for COVID-19. Diabetes, hypertension, and pneumonia were the most common diseases in both groups, but they were more frequent in the COVID-19 positive group ($p < 0.05$). Intubation, ICU admission, and hospital admission were also more common in the COVID-19 positive group, with rates of 6.78 %, 3.35 %, and 58.37 %, respectively ($p < 0.05$). Obesity, asthma, and immunosuppressive disorders were higher in the positive group ($p < 0.05$). However, COPD, smoking habits, and CVD were more common in the negative group ($p < 0.05$).

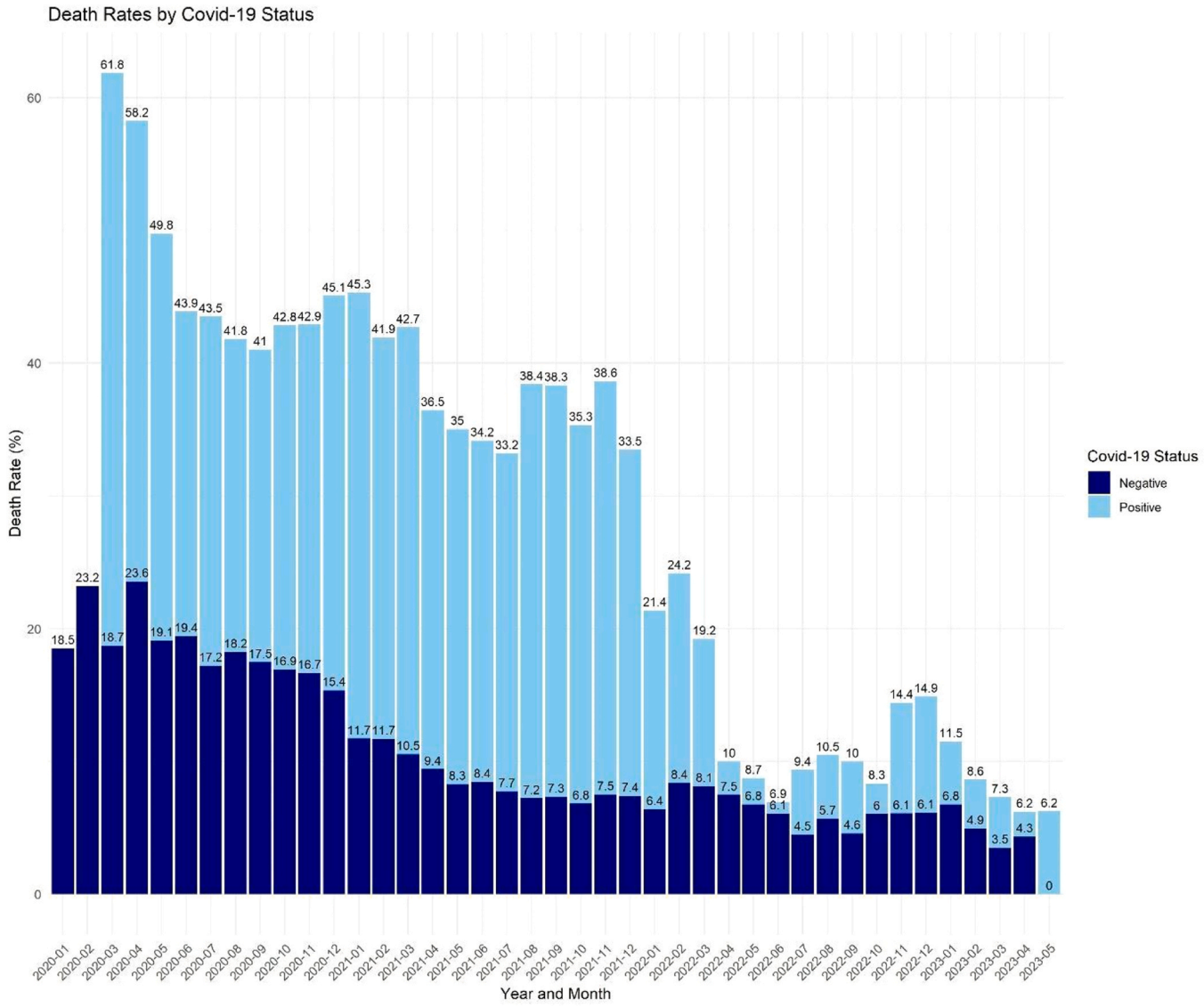

Fig. 2. The monthly mortality rates during the entire pandemic period.

Mortality rates

This study revealed an overall mortality rate of 19.33 % among CKD patients. Gender differences in mortality rates were observed: female CKD patients had a mortality rate of 17.3 %, while males had a higher rate of 21.64 %. Moreover, Fig. 2 presented the monthly mortality rates among COVID-19 positive and negative groups across the entire pandemic period. Counts of monthly deaths were shown in Figures E1 (in the Supplementary Materials).

Furthermore, Figures E2 (in the Supplementary Materials) showed distinct mortality rates among CKD patients with different comorbid conditions and smoking. The highest mortality rate of 47.86 % was due to CKD patients with pneumonia, while the asthma-CKD group had the lowest mortality rate of 11.56 %.

Mortality was significantly higher among COVID-19 positive patients compared to COVID-19 negative patients (33.9 % vs. 10.1 %, $p < 0.001$). We further evaluated subgroup mortality rates among CKD patients with and without COVID-19 (Figures E3 in the Supplementary Materials), revealing significant differences across various comorbidities. Pneumonia-CKD patients experienced the highest mortality rates, at 61.98 % with COVID-19 and 28.38 % without. Subsequently, diabetic-CKD patients had mortality rates of 43.31 % with COVID-19 compared to 13.17 % without COVID-19, while the asthma-CKD group had notably both the lowest mortality rates, at 19.08 % with COVID-19 and 6.27 % without. These findings underscore a significantly elevated mortality risk among CKD patients with COVID-19, with males consistently exhibiting higher mortality rates than females.

Machine learning models

An optimal XGBoost model was developed for predicting mortality. To improve the XGBoost models for predicting clinical outcome among renal patients, we performed hyperparameter tuning using 5-fold cross validation. For creating models, the datasets of each clinical outcome were divided into a training set (54,915 patients), a validation set (11,767 patients), and a test set (11,767 patients). These settings were then used to train the models on the entire training dataset. The performance of the trained XGBoost model was evaluated using the test datasets.

The XGBoost model for predicting mortality demonstrated superior performance on the held-out test set. It achieved an AUC of 0.77 (95 % CI: 0.76–0.77), accuracy of 0.71 (95 % CI: 0.70–0.72), precision of 0.80 (95 % CI: 0.79–0.81), recall of 0.72 (95 % CI: 0.71–0.73), F1-score of 0.75 (95 % CI: 0.75–0.76), specificity of 0.70 (95 % CI: 0.68–0.71), and negative predictive value of 0.59 (95 % CI: 0.57–0.60). Performance metrics were estimated with 1000 bootstrap iterations to provide confidence intervals, demonstrating stable predictive ability. Fig. 3 presents the ROC curve of the model for mortality prediction in the test set of CKD patients.

Feature importance

Fig. 4 shows the SHAP summary plot for the mortality prediction model. The top 10 most important features were *"COVID-19 status"*, *"Pneumonia"*, *"Intubation"*, *"Age_over_60"*, *"Sex"*, *"AgeLessThan20"*, *"ContactOtherCOVID"*, *"Age_21_30"*, *"Age_51_60"*, and *"Diabetes"*. The contributions of comorbidities to the mortality risk were ranked as *"Diabetes"*, *"CVD"*, *"Hypertension"*, *"Immunosuppression"*, *"Obesity"*, *"Asthma"*, and "*COPD*". Age again emerged as a dominant predictor.

In addition, Table 2 presents associations of each factor with mortality in CKD patients generated by the logistic regression. Several clinical conditions demonstrated statistically significant associations with increased odds of mortality: intubation (OR = 6.41, $p < 0.05$), pneumonia (OR = 1.89, $p < 0.05$), diabetes (OR = 1.07, $p < 0.05$), obesity (OR = 1.041, $p < 0.05$). Furthermore, advancing age was a significant predictor, with older age groups demonstrating a progressive increase in mortality risk compared to the reference group. These findings are consistent with those achieved by XGBoost and SHAP analysis.

Interactions of the factors

This study further explored the complex interactions of risk factors in mortality through advanced analytical techniques. By examining how different factors interact to associate with mortality, we revealed

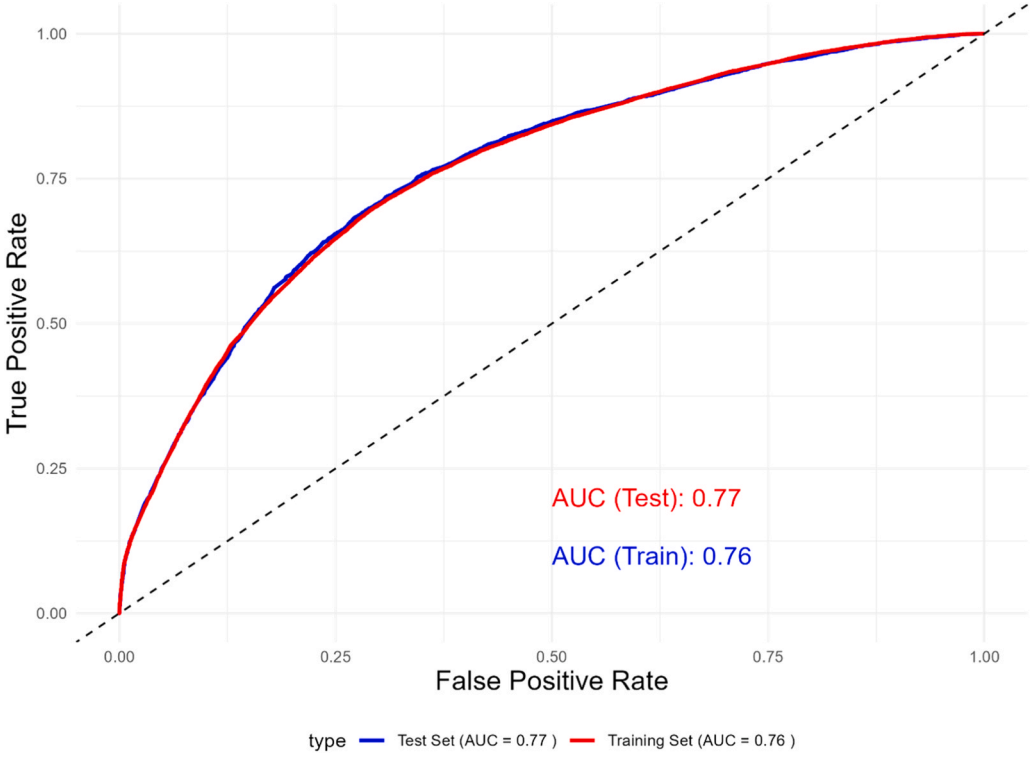

Fig. 3. ROC Curves for Mortality Prediction Model.

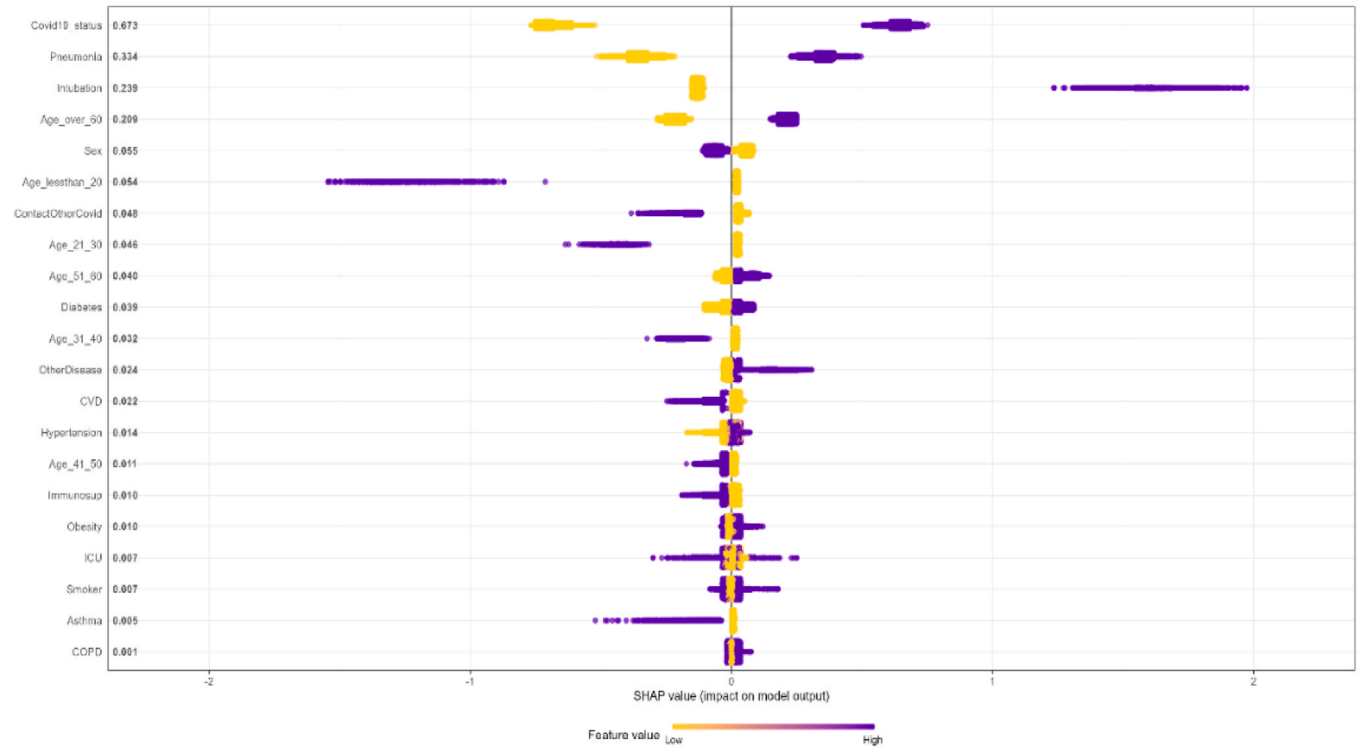

Fig. 4. SHAP Summary Plots for XGBoost Mortality Model: Each dot represents a SHAP value for a feature in a specific prediction. The position of a dot along the x-axis indicates the impact of feature on the model's prediction for that instance, with negative values on the left (decreasing the likelihood of outcome) and positive values on the right (increasing the likelihood of outcome). The numbers next to each feature on the y-axis represent the mean absolute SHAP value for that feature across all instances. The mean absolute SHAP value measures the overall importance of the feature, regardless of whether its effect is positive or negative.

Table 2
Multivariate logistic regression analysis of mortality predictors in CKD patients.

Variable	Odds Ratio (OR)	95 % CI (Lower)	95 % CI (Upper)	p-value
Sex	0.886714	0.857663	0.916738	1.49E-12
Intubation	6.407315	5.949022	6.905774	0
Pneumonia	1.892438	1.830789	1.956186	0
Diabetes	1.067424	1.028261	1.108098	0.000625
COPD	0.992393	0.927123	1.06205	0.825616
Asthma	0.842895	0.741303	0.957225	0.008759
Immunosuppression	0.92507	0.869039	0.984477	0.014366
Hypertension	1.006603	0.96757	1.04725	0.744432
Other Disease	1.242165	1.16433	1.325031	4.86E-11
CVD	0.896344	0.852681	0.942131	1.71E-05
Obesity	1.041298	0.996258	1.08831	0.072651
Smoker	1.028574	0.971324	1.089052	0.33435
Contact Other Covid	0.803805	0.766881	0.842417	8.00E-20
ICU	0.811194	0.742316	0.886324	3.72E-06
Covid19	3.856315	3.730326	3.986807	0
Age under 20	Ref	Ref	Ref	Ref
Age 21–30	2.097901	1.748281	2.529759	3.71E-15
Age 31–40	2.754542	2.312993	3.298297	4.17E-29
Age 41–50	3.577896	3.01208	4.27403	2.64E-46
Age 51–60	4.348655	3.66979	5.183133	1.41E-62
Age over - 60	5.803577	4.908237	6.903293	6.62E-91

nuanced insights about disease progression and patient outcomes. For instance, this study revealed four discrete zones corresponding to all possible combinations of infection status (Fig. 5A). Patients without either infection (pneumonia = 0, COVID-19 = 0) had near-zero SHAP values, indicating minimal contribution to predicting mortality. Presence of either pneumonia or COVID-19 alone was associated with modestly elevated SHAP values, suggesting a moderate independent increase in predicted mortality risk. Moreover, the presence of both pneumonia and COVID-19 (both = 1) exhibited markedly the highest SHAP interaction values, indicating a synergistic pattern in which concurrent infection substantially amplified the model's predicted mortality risk beyond the additive effects of each condition individually. Another case study of the interaction analysis between diabetes and age (Fig. 5B) demonstrated that the association of diabetes with mortality was not uniform across age groups, which warrants further research.

Figures S18-S28 (in Supplementary Materials) revealed intricate interactions among other comorbidities and demographic factors in predicting mortality.

Our findings highlighted that the impact of individual conditions can vary, with some combinations creating more synergistic effects on prediction of potential health outcomes. Local SHAP waterfall plot of CKD patients with diabetes revealed that within this subgroup, advanced age (>60 years) and diabetes status were the strongest positive contributors to mortality prediction (Figure S1 in Supplementary Materials). Covid-19 status showed a negative SHAP contribution in this subgroup, suggesting that once diabetes and CKD were present, Covid-19 infection did not substantially increase predicted risk beyond the high baseline conferred by these comorbidities. In the subgroup of CKD patients with obesity (Figure S7 in Supplementary Materials), SHAP waterfall analysis indicated that intubation status had the strongest positive association with mortality predictions (+0.0437), followed by obesity (+0.0266) and age > 60 years (+0.0258). Diabetes also contributed modestly to increased risk. For further observations, see Figures in the Supplementary Materials depicting waterfall plots for subgroup patients with different comorbidity profiles, demonstrating how each feature contributes to the predictions of mortality.

These results provided a comprehensive overview of the impact of various factors on mortality among renal patients, highlighting the complex interplay between demographic characteristics, comorbidities, and clinical outcomes. The machine learning models, combined with XAI techniques, offer additional insights into the relative importance of different factors and their interactions, potentially informing more

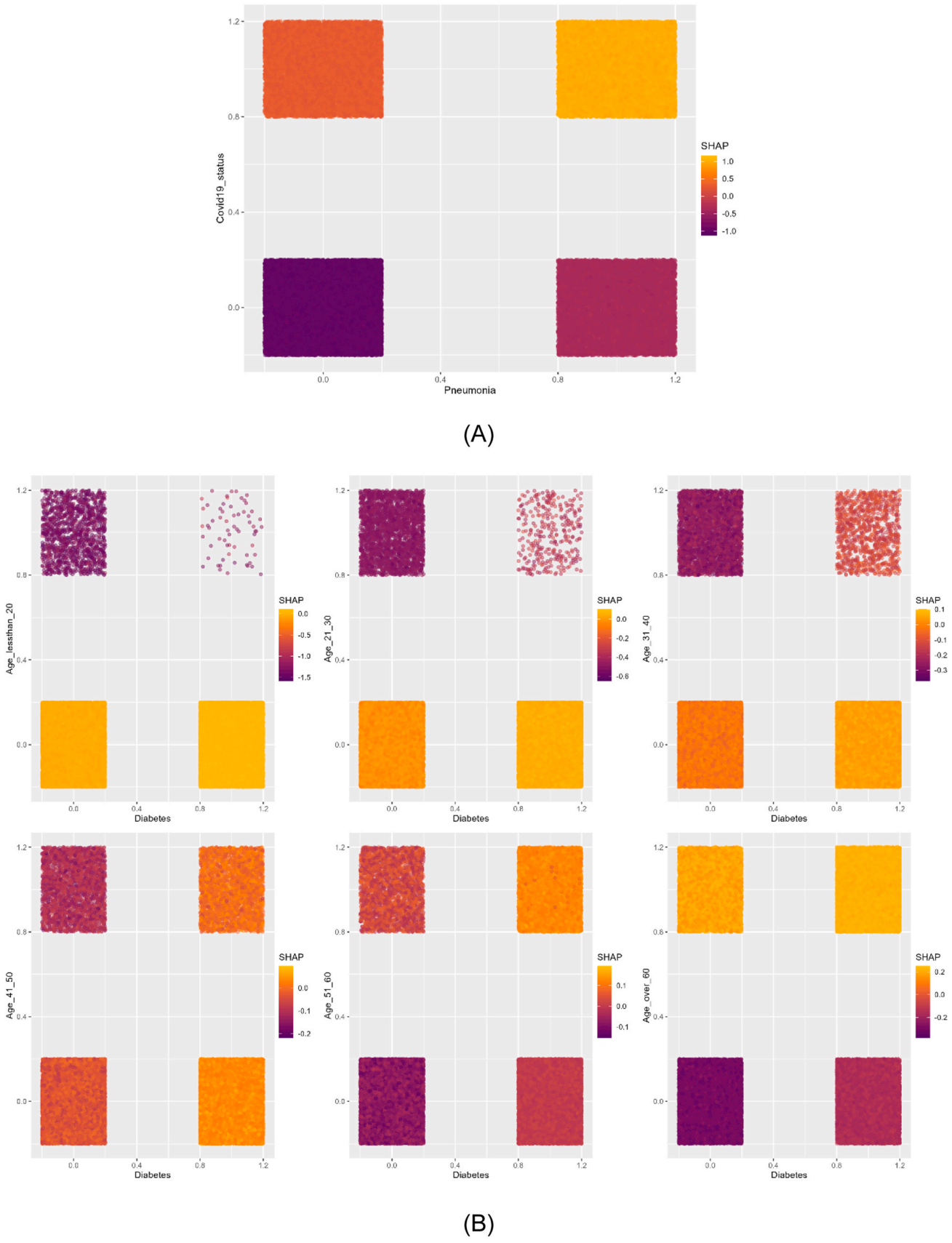

Fig. 5. Dependency Plots of Pneumonia-COVID-19 (A), and Diabetes-Age (B) in Mortality Prediction.

targeted risk assessment and management strategies for this high-risk population.

Discussion

Chronic kidney disease represents a significant public health concern, with effective risk stratification being crucial for timely intervention and management. This comprehensive study of 186,396 CKD patients in the national database during the COVID-19 pandemic has yielded several significant findings that contribute to our understanding of CKD outcomes and risk factors. Our findings highlight the complex interplay between CKD, comorbidities, and mortality, offering valuable insights for clinical management and risk stratification. These findings can be contextualized against international research, revealing both consistencies and important discrepancies.

Comparisons with other studies

In this study, the overall CKD mortality rate of 19.33 % and COVID-19-positive CKD mortality of 33.9 % aligns with global findings. An Italian study of CKD patients reported a crude mortality rate of 44.6 % among CKD patients with COVID-19 compared to 4.7 % in CKD patients without COVID-19 [7], which is actually higher than the Mexican findings. The UK's OpenSAFELY study of 17 million patients identified CKD stages 4–5 as conferring higher risk than diabetes or chronic heart disease, with dialysis showing an adjusted hazard ratio of 3.69 and CKD with eGFR $<$ 30 mL/min showing aHR of 2.52 [15].

This study found hypertension (64.39 %), diabetes (49.79 %), and obesity (16.46 %) as the most prevalent comorbidities. This pattern is generally consistent with international findings, though prevalence rates vary. An Italian CKD cohort reported 33.8 % had diabetes, 21.0 % had cardiovascular comorbidities, and 5.1 % had COPD [7]. The notably higher comorbidity rates in our CKD-specific cohort likely reflect the advanced disease burden in this population.

Our study found that COPD and smoking showed minimal mortality impact contrasts sharply with some international studies. A 2020 meta-analysis of 15 studies found COPD patients were at higher risk of severe COVID-19 disease, with a calculated risk ratio of 1.88 and associated mortality of 60 %, though COPD prevalence among COVID-19 cases was only 2 % [16]. Multiple international studies have documented the controversial "smoker's paradox," where current smokers appeared underrepresented among COVID-19 patients. Research identified several biases and knowledge gaps that may give the false impression smoking is protective in COVID-19, with claims of protective effects viewed as limited and questionable [17]. A Dutch study of 57,833 participants found current smokers were more likely to be non-responders to COVID-19 questionnaires, and these non-responders were more likely to have other established risk factors for SARS-CoV-2 infection, potentially explaining the paradox [18].

Our finding of higher male mortality (21.64 % vs. 17.3 % in females) in this study is consistent with global patterns. International reports generally identify older age, male sex, obesity, hypertension, diabetes, cardiovascular disease and chronic lung disease as risk factors for COVID-19 mortality [19].

Prevalence and impact of risk factors

Our analysis revealed a high prevalence of comorbidities among CKD patients during the COVID-19 pandemic, with diabetes, hypertension, and pneumonia being the most common. This aligns with previous studies before the pandemic [2]. The higher prevalence of these comorbidities in the COVID-19 positive group highlights the complex nature of CKD and its association with metabolic disorders, suggesting that these conditions may increase susceptibility to SARS-CoV-2 infection or lead to more frequent testing and diagnosis.

Fig. 2 showed a general trend of the drops of mortality rates in both COVID-19 positive and negative groups. Such a declining trend could be due to increasing numbers of participants being tested (i.e. the denominators) along with the pandemic and gradually advancing measures of tackling the SARS-CoV-2 infections. The strong association between pneumonia and both COVID-19 infection and mortality underscores the importance of respiratory health in renal patients, especially given its emergence as a major risk factor for adverse outcomes in our predictive models. This finding is consistent with the previous research indicating that CKD patients are at increased risk for respiratory infections and related complications [7]. Such strong links between pneumonia and adverse outcomes emphasize the need for vigilant monitoring and aggressive management of respiratory symptoms in this vulnerable population.

Diabetes significantly increases the risk of mortality in CKD patients, consistent with findings from other studies. The risk of death is particularly high for individuals who have both diabetes and CKD, compared to the general population [20]. For people with type 2 diabetes and diabetic kidney disease (DKD), the risk remains alarmingly high, with one out of every 27 individuals with DKD dying each year [21](p2). Another recent study found that CKD may affect 25–50 % of people with diabetes, leading to higher mortality rates and increased healthcare costs associated with DKD [22].

Notably, COPD and smoking habits did not show strong contributions to clinical outcomes in this study. COPD has been reported as an associated risk factor with increased odds of mortality during the pandemic [23]. Many meta-analyses of observational studies confirm that smoking is linked to severe COVID-19 and poor outcomes [24–26], due to increased levels of angiotensin-converting-enzyme-2 (ACE2), immune system disturbances, blood vessel damage, and a higher likelihood of blood clots [26,27]. Our results fall in the middle of the risk spectrum, showing neither a protective nor harmful effect, which warrants careful consideration of potential explanations. In CKD patients, the overwhelming burden of renal disease and associated comorbidities may overshadow the independent effects of COPD and smoking. When multiple strong risk factors coexist, machine learning models may attribute mortality risk primarily to the most dominant predictors, potentially masking weaker associations.COPD prevalence was only 5.9 % and smoking prevalence was 9.53 % in this CKD population, which is notably lower than in many general COVID-19 cohorts. This lower prevalence could limit statistical power to detect associations. Additionally, the study notes that COPD and smoking were actually more common in the COVID-19 negative group, which runs counter to expected patterns and suggests possible selection bias or survival bias effects. CKD patients with severe COPD may have been underrepresented in this dataset due to healthcare avoidance that severely ill patients may have avoided hospitals during the pandemic; or protective medications that CKD patients with respiratory comorbidities may have been on protective treatments (inhaled corticosteroids, bronchodilators) that modified COVID-19 outcomes. The so-called "smoker's paradox" has also been revealed in acute myocardial infarction [28]. The "smoker's paradox" in COVID-19 literature may reflect collider bias that COVID-19 testing was not random while smokers who were tested may represent a selected subgroup; also reflect age confounding that younger smokers might have been more likely to survive despite smoking. Geographic and healthcare system factors specific to Mexico may be relevant, such as different smoking prevalence patterns compared to Western countries, and cultural factors affecting smoking disclosure. This topic merits further investigations.

Regarding mortality of CKD patients, a major public health issue, our findings indicated that clinical interventions, such as intubation, are significantly associated with an increased likelihood of death. While Jang et al. observed that prolonged intubation decreases the survival rates of patients receiving mechanical ventilation [29], the specific impact of intubation on mortality among CKD patients remains unclear and need to be more studied by future research. Interestingly, admission to the ICU appears to be linked to a lower risk of mortality, potentially

highlighting the effectiveness of intensive care in managing severe conditions and improving survival outcomes despite the critical status of these patients [30].

Not surprisingly, COVID-19 infection emerged as one of the strongest risk factors for death in patients with CKD. Gilbertoni found that COVID-19-related mortality in CKD patients was about ten times higher than in those without COVID-19 [7]. Therefore, these individuals with COVID-19 need careful monitoring and management to reduce their risk of death [31].

In this study, the 10.1 % mortality rate observed in COVID-negative CKD patients, while substantially lower than the 33.9 % rate in COVID-positive patients, remains considerable and warrants careful interpretation. Several factors may contribute to this finding. First CKD patients commonly experience complications that can phenotypically resemble COVID-19, including pulmonary edema and fluid overload presenting with dyspnea and respiratory distress; uremic encephalopathy causing altered mental status similar to COVID-19-related neurological symptoms [32]. These overlapping clinical presentations may have led to COVID-19 testing in patients whose symptoms were primarily driven by CKD-related complications rather than viral infection. This could explain a portion of the mortality observed in the COVID-negative group, as these patients were genuinely ill from non-COVID causes that prompted testing.

Additionally, RT-PCR testing for SARS-CoV-2, while considered the gold standard, has known limitations with sensitivity ranging from 70 % to 95 % depending on timing of testing relative to symptom onset, specimen quality, and viral load [33,34]. False negative rates are particularly high when testing occurs very early or late in the disease course, and multiple studies have documented that a single negative RT-PCR does not completely exclude COVID-19 infection [35]. It is plausible that some patients in our COVID-negative cohort had false negative test results and were actually infected with SARS-CoV-2. A meta-analysis by Kucirka et al. demonstrated that the false-negative rate of RT-PCR varies from 100 % on day 1 of infection to 20 % on day 8 (3 days post-symptom onset), increasing again as the infection progresses [36]. This would be expected to partially attenuate the observed mortality difference between groups (making COVID-negative mortality appear higher than it truly should be).

COVID-19 underdiagnosis was more likely during early pandemic phases (early 2020) when testing availability was limited, clinical awareness of atypical COVID-19 presentations was still developing, and testing criteria were more restrictive [37]. Our dataset spans from January 2020 to May 2023, encompassing these early periods of limited testing capacity. However, several factors mitigate this concern. The dataset includes the entire pandemic period, with the majority of cases occurring after mid-2020 when testing became more widely available. Mexico implemented broad testing strategies for suspected respiratory infections, particularly in high-risk populations including CKD patients [8]. Our case definition (Section 2.2) required RT-PCR confirmation, representing the most specific diagnostic approach available. It is important to contextualize the 10.1 % mortality rate within the known high baseline mortality risk of CKD patients even in the absence of COVID-19. Pre-pandemic studies have documented annual mortality rates of 6–8 % in stage 3–5 CKD patients [38,39], with substantially elevated mortality risk from cardiovascular disease (present in 11.65 % of our cohort), infections, and CKD progression. Our study period spanned 3.25 years (January 2020 - May 2023), during which non-COVID mortality would naturally accumulate. The 10.1 % mortality rate in COVID-negative patients may largely reflect the underlying disease burden of this high-risk population rather than indicating widespread COVID-19 underdiagnosis.

While some degree of COVID-19 misclassification (both false negatives and CKD complications mimicking COVID-19) is possible, several factors suggest this does not substantially compromise our key findings. Our primary aim was to identify risk factors for mortality in CKD patients during the pandemic period and develop predictive models—both objectives remain valid regardless of some potential misclassification. The very large sample size (178,462 patients with confirmed status) provides robustness against moderate levels of misclassification [40]. Our XAI-based models successfully identified pneumonia as the major predictor of mortality in COVID-positive patients (Fig. 4), demonstrating clinical validity

Health disparities in gender and age

Our analysis revealed that males with renal disorders are more likely to experience adverse clinical outcomes compared to females during the pandemic. This finding aligns with research suggesting that males are at a higher risk for severe health outcomes before the pandemic [41,42]. The exact reasons behind these gender differences are not entirely clear, but they may be linked to the higher prevalence of kidney disease risk factors and more rapid progression of CKD in men [43]. These observations highlight the need for further exploration of how gender influences the development and progression of CKD.

The COVID-19 infection rate of 39.9 % among CKD patients is alarming and suggests a heightened susceptibility to infection in this population. The higher infection rate observed in male patients (41.2 % vs. 38.4 % in females) aligns with global trends of COVID-19 disproportionately affecting men. This gender disparity in infection rates may be attributed to biological factors, such as differences in immune responses, or behavioural factors like adherence to preventive measures.

When comparing different age groups, all older age groups have a significantly higher likelihood of COVID-19 infection and mortality compared to those under 20, underscores the increasing burden of CKD in aging populations, a trend observed globally. Previous studies on COVID-19 and mortality rate have consistently linked age to adverse clinical outcomes [44,45].

Insights into subpopulations

The use of SHAP values for model interpretation provided valuable insights into the relative importance of different factors in predicting outcomes as shown in the summary plot (Fig. 4). Interestingly, SHAP analysis further offered a way of examining the impact of risk factors on clinical outcomes for a subpopulation, such as subgroups of individuals with specific comorbidities via waterfall plots. As shown in Figures S1-S17 (Supplementary Materials), we generated waterfall plots for various subgroups of individuals with specific comorbidities to gain clearer and more informative insights into the mortality prediction in different subpopulation of CKD patients with specific pre-existing disorders. For example, in subgroup analysis restricted to CKD patients with diabetes (Figure S1, Supplementary Materials) underscores the major role of acute respiratory compromise, as evidenced by intubation being the most influential feature in mortality prediction. This aligns with clinical evidence that obesity predisposes to severe respiratory failure and poor ventilatory reserve. Age $>$ 60 years and obesity itself also independently increased predicted risk, consistent with their known biological roles in adverse outcomes.

These findings emphasize that in comorbidity-heavy populations, local model explanations can reveal nuanced interactions, where dominant baseline risks (such as obesity or diabetes) may overshadow the incremental effects of acute infections like Covid-19. Taken together, these findings demonstrate that machine learning interpretability methods such as SHAP can provide clinically meaningful insights, revealing both global drivers of risk and subgroup-specific interactions. However, It is important to note that SHAP values waterfall plots provide local explanations that are conditional on the specific subgroup and the trained model, and should not be interpreted as direct causal effects. Apparent counterintuitive findings, such as Covid-19 status contributing negatively to mortality predictions in certain subgroups, reflect how the model assigns relative importance based on the data distribution and feature interactions, rather than true protective effects. These results

may be influenced by subgroup composition or sample imbalance. Therefore, local explanations should be interpreted in conjunction with global SHAP summary plots, which provide an overall picture of feature importance across the full cohort. Considering both perspectives allows a more balanced understanding of the global drivers of mortality risk as well as subgroup-specific variations.

The SHAP interaction analysis identified a pronounced synergistic effect between pneumonia and COVID-19 on mortality risk prediction. Patients with both infections had a disproportionately higher model-predicted risk compared to those with only one infection, implying that the model detected a nonlinear interaction between these respiratory conditions. This finding aligns with clinical evidence showing that bacterial or secondary pneumonia in COVID-19 patients is associated with severe respiratory failure, prolonged hospitalization, and increased mortality. Although SHAP values reflect model behavior rather than causal inference, this result supports the clinical understanding that co-infection exacerbates disease severity and provides model-based evidence for the heightened vulnerability of these patients.

The observed interaction between diabetes and age provides important clinical insight into the age-dependent influence of metabolic disease on mortality among CKD patients. The stronger effect of diabetes on mortality prediction in younger individuals suggests that diabetes confers a disproportionate relative risk early in life, when competing age-related risks are otherwise low. This pattern aligns with epidemiological evidence showing that diabetes-related mortality risk is most pronounced in younger adults with CKD, where early metabolic derangements and microvascular complications may accelerate disease progression. In contrast, the diminishing incremental contribution of diabetes with advancing age indicates that, in older CKD patients, baseline vulnerability associated with aging and multimorbidity dominates overall mortality risk. The smaller separation of SHAP values between diabetic and non-diabetic elderly patients implies that, beyond a certain age threshold, diabetes adds relatively less discriminative information to the model's mortality prediction. Importantly, the model also revealed a synergistic effect when diabetes and advanced age co-occurred, substantially increasing the predicted probability of mortality. This finding highlights the clinical importance of integrated risk management in older diabetic CKD patients, where the convergence of metabolic, vascular, and age-related factors likely compounds vulnerability.

This study highlights the potential of explainable machine learning approaches, particularly SHAP-based interpretation, to elucidate complex and nonlinear interactions between comorbidities in CKD patients, offering a more nuanced understanding of mortality risk beyond traditional statistical modeling.

Translating insights into actionable clinical strategies

The insights from this study can be translated into actionable clinical strategies that go beyond merely confirming well-established risk factors. These strategies focus on leveraging the study's findings to refine clinical interventions, risk stratification, and healthcare resource allocation, particularly during pandemics or other public health crises.

Traditional CKD management does not consider the dynamic interplay of multiple comorbidities in real-time. The high prevalence of hypertension (64.39 %) and diabetes (49.79 %) underscores the need for aggressive management of these conditions in CKD patients. This includes tighter blood pressure control (<130/80 mmHg) and optimized glycemic management (e.g., HbA1c $\leq$7 %) during periods of increased risk (e.g., flu season, pandemic surges) to reduce mortality risks. Obesity (16.46 %) was associated with worse outcomes, highlighting the need for weight management programs, including dietary interventions and pharmacologic therapies like GLP-1 receptor agonists. Going beyond general CKD management, targeted management protocols for CKD patients with these specific comorbidities also involve collaboration between nephrologists, endocrinologists, cardiologists, and dieticians to provide comprehensive care, and patient education with enhanced education programs focusing on self-management of comorbidities, including diet, exercise, and medication adherence.

CVD was present in 11.65 % of CKD patients and more common in COVID-19-negative patients, suggesting a need for routine cardiovascular screening for CVD risk factors (hypertension, hyperlipidemia) in CKD patients, especially those who were COVID-19 negative during the pandemic. CKD patients with CVD can also benefit from lifestyle modifications, including a balanced diet, regular exercise, and smoking cessation. Proactive CVD management strategies also need medication optimisation to manage CVD risk factors, such as statins for hyperlipidemia and antihypertensive medications for hypertension.

Pneumonia was identified as the major predictor of mortality in COVID-19-positive CKD patients, yet it is preventable with early intervention. This calls for enhanced vaccination strategies, including pneumococcal and influenza vaccines, as well as early use of antiviral or antibiotic therapies to prevent and aggressively manage respiratory infections. CKD patients often receive suboptimal respiratory monitoring compared to other high-risk groups. Regular respiratory monitoring, such as pulse oximetry at home, can help detect early signs of respiratory distress in high-risk patients. Proactive measures of prevention and management of respiratory infection can also benefit from rapid diagnostic testing for respiratory pathogens (including influenza, RSV, and COVID-19) in CKD patients presenting with respiratory symptoms, and timely antiviral or antibiotic therapy based on the identified pathogen, considering potential renal toxicity of medications.

Given the higher rates of intubation (6.78 %) and ICU admission (3.35 %) among COVID-19-positive CKD patients, hospitals should prioritize respiratory support training for healthcare staff and ensure adequate availability of ventilators during surges.

Males CKD patients had a higher COVID-19 positivity rate (41.2 %) and mortality rate (21.64 %). Most clinical protocols treat males and females the same, despite clear evidence of sex differences in immune responses and disease progression. Gender-specific strategies, such as targeted awareness campaigns, tailored education, smoking cessation, weight management etc need to be implemented to address these disparities.

With 42 % of the cohort aged over 60 years experiencing disproportionately high mortality rates, older CKD patients should receive priority access to vaccinations, antiviral therapies, and telemedicine consultations.

For immunosuppressed CKD patients (e.g., transplant recipients), alternative vaccine administration methods (e.g., intradermal delivery) could improve immune responses.

This study underscores that comorbidities in CKD patients don't act in isolation—they interact in complex ways that AI-driven healthcare models can help unravel. These actionable strategies move beyond simply confirming risk factors to implementing innovative, proactive, and personalized solutions that can directly reduce mortality rates in CKD patients, especially during times of increased risk such as pandemics or seasonal epidemics.

Strengths and limitations

This study is innovative in several ways. Firstly, this study specifically targeted CKD patients, a high-risk group during the COVID-19 pandemic, providing valuable insights for their clinical management. Moreover, this study analysed data from 186,396 CKD patients, providing substantial analytics power and increasing the reliability of the findings. Thirdly, we examined multiple comorbidities and their impact on mortality from national healthcare data covering the entire COVID-19 pandemic, offering a holistic view of risk factors in CKD patients. Fourthly, this study revealed differences in outcomes between male and female patients, contributing to our understanding of gender disparities in CKD and COVID-19. Fifthly, this study incorporated XAI techniques to enhance the clinical relevance of the findings. Finally, the

findings have direct implications for patient care, risk stratification, and clinical decision-making in the management of renal patients, especially in the context of pandemics or other seasonal periods for infectious diseases.

While our study provided valuable insights, several limitations should be acknowledged. Firstly, the retrospective nature of our analysis may limit the ability to establish causal relationships and control for all potential confounding factors. Secondly, there may be selection bias in COVID-19 reporting and testing, particularly early in the pandemic. This could affect the accuracy of COVID-19 prevalence estimates and associated risk factors. Thirdly, this study focused on specific comorbidities and demographic factors. Other potentially important variables, such as socioeconomic factors or access to healthcare, were not accounted for in the analysis. Fourthly, the CKD and comorbidities in this data were collected by participant's self-reports before their COVID-19 testing. Although such self-reports of the underlining conditions were largely credible, the severity of these conditions or degree of recovery from these diseases cannot be accurately reflected. In addition, this national database does not include detailed pharmacological treatment information and different variants and treatment protocols (including vaccinations). The dataset was primarily designed for epidemiological surveillance of respiratory diseases and focuses on demographic characteristics, comorbidities, clinical presentations, and outcomes rather than therapeutic interventions. This is a common limitation of large-scale public health surveillance databases, which prioritize breadth of coverage over depth of clinical detail [46]. Sixthly, this study used data from Mexico's national healthcare system, reflecting Mexican healthcare practices and policies. However, the findings of this study should be generalisable to the populations with similar healthcare settings.

Conclusions

Our study highlights the importance of managing comorbid conditions and respiratory infections in renal patients during the COVID-19 pandemic. The use of advanced machine learning models with XAI techniques offered a comprehensive approach to understanding and predicting risk in this vulnerable population. As the global burden of kidney disease continues to grow, these insights will inform development of targeted strategies to improve patient care and outcomes during pandemics or other seasonal periods for infectious diseases.

Ethics statement

This study used a publicly available dataset to analyse clinical outcomes and develop predictive machine learning models. The dataset was anonymized and followed all relevant data protection regulations, ensuring that no personal information was included.

CRediT authorship contribution statement

Shang-Ming Zhou: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. **Zeinab Abdollahi:** Formal analysis, Investigation, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Donald Fraser:** Validation, Writing – review & editing. **Lin Huo:** Funding acquisition, Validation, Writing – review & editing.

Declaration of Competing Interest

We declare no conflicts of interest.

Acknowledgements

SMZ was supported in part by the UK CDT in Artificial Intelligence, Machine Learning and Advanced Computing (EP/S023992/1) and the international collaboration with Guangxi University "Digital ASEAN Cloud Big Data Security and Mining Technology" Innovation Team. LH was supported in part by the Major Project of National Social Science Foundation of China (16ZDA0092), and in part by Guangxi University "Digital ASEAN Cloud Big Data Security and Mining Technology" Innovation. The funders had no involvement in the study design, collection, analysis and interpretation of data, writing of the report and decision to submit the article for publication.

Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.annepidem.2025.11.005.

Data availability

The data used for this study covering the entire COVID-19 pandemic is available from national healthcare database - General Directorate of Epidemiology in Mexico. The scripts of this study will be shared upon request.

References

[1] Evans M, Lewis RD, Morgan AR, et al. A narrative review of chronic kidney disease in clinical practice: current challenges and future perspectives. Adv Ther 2022;39 (1):33–43. https://doi.org/10.1007/s12325-021-01927-z.
[2] Jager KJ, Kovesdy C, Langham R, Rosenberg M, Jha V, Zoccali C. A single number for advocacy and communication—worldwide more than 850 million individuals have kidney diseases. Nephrol Dial Transpl 2019;34(11):1803–5. https://doi.org/10.1093/ndt/gfz174.
[3] Francis A, Harhay MN, Ong ACM, et al. Chronic kidney disease and the global public health agenda: an international consensus. Nat Rev Nephrol 2024;20(7): 473–85. https://doi.org/10.1038/s41581-024-00820-6.
[4] Russell CD, Lone NI, Baillie JK. Comorbidities, multimorbidity and COVID-19. Nat Med 2023;29(2):334–43. https://doi.org/10.1038/s41591-022-02156-9.
[5] Ong KL, Stafford LK, McLaughlin SA, et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2023;402(10397): 203–34. https://doi.org/10.1016/S0140-6736(23)01301-6.
[6] Foreman KJ, Marquez N, Dolgert A, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. Lancet 2018;392(10159):2052–90. https://doi.org/10.1016/S0140-6736(18)31694-5.
[7] Gibertoni D, Reno C, Rucci P, et al. COVID-19 incidence and mortality in non-dialysis chronic kidney disease patients. In: Remuzzi G, editor. PLOS ONE, 16; 2021, e0254525. https://doi.org/10.1371/journal.pone.0254525.
[8] Secretaría de Salud. Datos Abiertos - Dirección General de Epidemiología | Secretaría de Salud | Gobierno | gob.mx. Gob. 2020. Accessed October 28, 2024. ⟨https://www.gob.mx/salud/documentos/datos-abiertos-152127⟩.
[9] Statement on the Second Meeting of the International Health Regulations (2005) Emergency Committee Regarding the Outbreak of Novel Coronavirus (2019-nCoV). Accessed October 28, 2024. ⟨https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations⟩-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).
[10] World Health Organization. Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. Accessed May 19, 2023. ⟨https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations⟩⟨-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic⟩.
[11] Montomoli J, Romeo L, Moccia S, et al. Machine learning using the extreme gradient boosting (XGBoost) algorithm predicts 5-day delta of SOFA score at ICU admission in COVID-19 patients. J Intensive Med 2021;1(2):110–6. https://doi.org/10.1016/j.jointm.2021.09.002.
[12] Hakkal S, Lahcen AA. XGBoost to enhance learner performance prediction. Comput Educ Artif Intell 2024;7:100254. https://doi.org/10.1016/j.caeai.2024.100254.
[13] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 2017.
[14] Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion 2020;58:82–115. https://doi.org/10.1016/j.inffus.2019.12.012.
[15] Council ERA-EDTA, Ortiz A, Cozzolino M, et al. Chronic kidney disease is a key risk factor for severe COVID-19: a call to action by the ERA-EDTA. Nephrol Dial Transpl 2021;36(1):87–94. https://doi.org/10.1093/ndt/gfaa314.
[16] Alqahtani JS, Oyelade T, Aldhahir AM, et al. Prevalence, severity and mortality associated with COPD and smoking in patients with COVID-19: A rapid systematic review and meta-analysis. In: Bhatt GC, editor. PLOS ONE, 15; 2020, e0233147. https://doi.org/10.1371/journal.pone.0233147.

[17] Usman MS, Siddiqi TJ, Khan MS, et al. Is there a smoker's paradox in COVID-19? BMJ EvidBased Med 2021;26(6):279–84. https://doi.org/10.1136/bmjebm-2020-111492.
[18] Kramer I, Zhu Y, Van Westen-Lagerweij NA, Dekker LH, Mierau JO, Croes EA. New insights into the paradox between smoking and the risk of SARS-CoV-2 infection (COVID-19): insufficient evidence for a causal association. Scand J Public Health 2025;53(5):552–9. https://doi.org/10.1177/14034948241253690.
[19] Gansevoort RT, Hilbrands LB. CKD is a key risk factor for COVID-19 mortality. Nat Rev Nephrol 2020;16(12):705–6. https://doi.org/10.1038/s41581-020-00349-4.
[20] Afkarian M, Sachs MC, Kestenbaum B, et al. Kidney disease and increased mortality risk in type 2 diabetes. J Am Soc Nephrol 2013;24(2):302–8. https://doi.org/10.1681/ASN.2012070718.
[21] González-Pérez A, Saez M, Vizcaya D, Lind M, Garcia Rodriguez L. Incidence and risk factors for mortality and end-stage renal disease in people with type 2 diabetes and diabetic kidney disease: a population-based cohort study in the UK. BMJ Open Diabetes Res Care 2021;9(1):e002146. https://doi.org/10.1136/bmjdrc-2021-002146.
[22] Jairoun AA, Ping CC, Ibrahim B. Predictors of chronic kidney disease survival in type 2 diabetes: a 12-year retrospective cohort study utilizing estimated glomerular filtration rate. Sci Rep 2024;14(1):9014. https://doi.org/10.1038/s41598-024-58574-x.
[23] Gerayeli FV, Milne S, Cheung C, et al. COPD and the risk of poor outcomes in COVID-19: a systematic review and meta-analysis. EClinicalMedicine 2021;33:100789. https://doi.org/10.1016/j.eclinm.2021.100789.
[24] Del Sole F, Farcomeni A, Loffredo L, et al. Features of severe COVID-19: a systematic review and meta-analysis. Eur J Clin Invest 2020;50(10):e13378. https://doi.org/10.1111/eci.13378.
[25] Zheng Z, Peng F, Xu B, et al. Risk factors of critical & mortal COVID-19 cases: a systematic literature review and meta-analysis. J Infect 2020;81(2):e16–25. https://doi.org/10.1016/j.jinf.2020.04.021.
[26] Salah HM, Sharma T, Mehta J. Smoking doubles the mortality risk in COVID-19: a meta-analysis of recent reports and potential mechanisms. Published online October 7, Cureus 2020. https://doi.org/10.7759/cureus.10837.
[27] Cai G, Bossé Y, Xiao F, Kheradmand F, Amos CI. Tobacco smoking increases the lung gene expression of ACE2, the receptor of SARS-CoV-2. Am J Respir Crit Care Med 2020;201(12):1557–9. https://doi.org/10.1164/rccm.202003-0693LE.
[28] Paradossi U, De Caterina AR, Trimarchi G, et al. The enigma of the 'smoker's paradox': results from a single-center registry of patients with STEMI undergoing primary percutaneous coronary intervention. Cardiovasc Revasc Med 2024;69:42–9. https://doi.org/10.1016/j.carrev.2024.06.007.
[29] Jang CS, Wang JD. Predicting mortality and life expectancy in patients under prolonged mechanical ventilation and maintenance dialysis. J Palliat Med 2020;23(1):74–81. https://doi.org/10.1089/jpm.2018.0646.
[30] Hittesdorf E, Panzer O, Wang D, et al. Mortality and renal outcomes of patients with severe COVID-19 treated in a provisional intensive care unit. J Crit Care 2021;62:172–5. https://doi.org/10.1016/j.jcrc.2020.12.012.
[31] Cai R, Zhang J, Zhu Y, Liu L, Liu Y, He Q. Mortality in chronic kidney disease patients with COVID-19: a systematic review and meta-analysis. Int Urol Nephrol 2021;53(8):1623–9. https://doi.org/10.1007/s11255-020-02740-3.
[32] Brouns R, De Deyn PP. Neurological complications in renal failure: a review. Clin Neurol Neurosurg 2004;107(1):1–16. https://doi.org/10.1016/j.clineuro.2004.07.012.
[33] Woloshin S, Patel N, Kesselheim AS. False negative tests for SARS-CoV-2 infection — challenges and implications. N Engl J Med 2020;383(6). https://doi.org/10.1056/NEJMp2015897.
[34] Fang FC, Naccache SN, Greninger AL. The laboratory diagnosis of coronavirus disease 2019— frequently asked questions. Clin Infect Dis 2020;71(11):2996–3001. https://doi.org/10.1093/cid/ciaa742.
[35] West CP, Montori VM, Sampathkumar P. COVID-19 testing: the threat of false-negative results. Mayo Clin Proc 2020;95(6):1127–9. https://doi.org/10.1016/j.mayocp.2020.04.004.
[36] Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure. Ann Intern Med 2020;173(4):262–7. https://doi.org/10.7326/M20-1495.
[37] Patel R, Babady E, Theel ES, et al. Report from the American Society for Microbiology COVID-19 International Summit, 23 March 2020: value of diagnostic testing for SARS–CoV-2/COVID-19. mBio 2020;11(2):e00722-20. https://doi.org/10.1128/mBio.00722-20.
[38] Ang YG, Heng BH, Saxena N, Liew STA, Chong PN. Annual all-cause mortality rate for patients with diabetic kidney disease in Singapore. J Clin Transl Endocrinol 2016;4:1–6. https://doi.org/10.1016/j.jcte.2016.01.002.
[39] Go AS, Chertow GM, Fan D, McCulloch CE, Hsu C yuan. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med 2004;351(13):1296–305. https://doi.org/10.1056/NEJMoa041031.
[40] Rothman KJ, Lash TL, Greenland S. Modern Epidemiology. Lippincott Williams & Wilkins; 2012.
[41] Ricardo AC, Yang W, Sha D, et al. Sex-related disparities in CKD progression. J Am Soc Nephrol 2019;30(1):137–46. https://doi.org/10.1681/ASN.2018030296.
[42] Hockham C, Schanschieff F, Woodward M. Sex differences in CKD-associated mortality from 1990 to 2019: data from the global burden of disease study. Kidney Med 2022;4(10):100535. https://doi.org/10.1016/j.xkme.2022.100535.
[43] García GG, Iyengar A, Kaze F, Kierans C, Padilla-Altamira C, Luyckx VA. Sex and gender differences in chronic kidney disease and access to care around the globe. Semin Nephrol 2022;42(2):101–13. https://doi.org/10.1016/j.semnephrol.2022.04.001.
[44] Grasselli G, Zangrillo A, Zanella A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy. JAMA 2020;323(16):1574. https://doi.org/10.1001/jama.2020.5394.
[45] Guan W jie, Liang W hua, Zhao Y, et al. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. Eur Respir J 2020;55(5). https://doi.org/10.1183/13993003.00547-2020. 2000547.
[46] Denaxas S, Gonzalez-Izquierdo A, Direk K, et al. UK phenomics platform for developing and validating electronic health record phenotypes: CALIBER. J Am Med Inf Assoc 2019;26(12):1545–59. https://doi.org/10.1093/jamia/ocz105.