공학박사 학위논문

Assessing risks of acute kidney injury and glomerulonephritis in the smart healthcare era

– A convergent approach using hospital time–series data and machine learning –

스마트 헬스케어 시대의 급성 신손상과 사구체신염 위험도 평가

– 병원 시계열 데이터와 머신러닝의 융합적 접근 –

2025년 02월

서울대학교 융합과학기술대학원

헬스케어융합학과

이 진 영

Assessing risks of acute kidney injury and glomerulonephritis in the smart healthcare era

– A convergent approach using hospital time–series data and machine learning –

지도 교수 곽 노 준
공동지도 교수 김 세 중

이 논문을 공학박사 학위논문으로 제출함
2025년 02월

서울대학교 대학원
융합과학기술대학원 헬스케어융합학과
이 진 영

이진영의 공학박사 학위논문을 인준함

2025년 02월

위 원 장 _____ 류 주 석 _____ (인)

부위원장 _____ 김 세 중 _____ (인)

위 원 _____ 곽 노 준 _____ (인)

위 원 _____ 백 선 하 _____ (인)

위 원 _____ 이 상 철 _____ (인)

Abstract

Kidney disease, encompassing acute and chronic conditions, is a major challenge for modern medicine and healthcare systems around the world. The effective management of these diseases requires both an understanding of the multiple risk factors and the development of robust predictive strategies, as a variety of underlying pathologies may contribute to the decline of renal function over time. The importance of time series data in nephrology is increasingly recognized. Time series data play a critical role in tracking disease progression and implementing prevention strategies. In-depth analysis of time series data using advanced analytical techniques and artificial intelligence offers innovative approaches to the prediction and monitoring of kidney disease, potentially improving the quality of patient care and the efficiency of healthcare systems. This research combines two studies: one investigating environmental factors affecting renal function in primary glomerulonephritis, and another applying machine learning in acute kidney injury prediction, providing a comprehensive view of data-driven approaches in different types of kidney diseases.

(1) Air quality and kidney health: Assessing the effects of PM_{10}, $PM_{2.5}$, CO, and NO_2 on renal function in primary glomerulonephritis

Background: While extensive studies have elucidated the relationships between exposure to air pollution and chronic diseases, such as cardiovascular disorders and diabetes, the intricate effects on specific kidney diseases, notably primary glomerulonephritis (GN)—an immune-mediated kidney ailment—are less well understood.

Considering the escalating incidence of GN and notable gap in research on its association with air quality, investigation is dedicated to examining the long-term effects of air pollutants on renal function in individuals diagnosed with primary GN.

Methods: This retrospective cohort analysis was conducted on 1394 primary GN patients who were diagnosed at Seoul National University Bundang Hospital and Seoul National University Hospital. Utilizing time-varying Cox regression and linear mixed models (LMM), we examined the effect of yearly average air pollution levels on renal function deterioration (RFD) and change in estimated glomerular filtration rate (eGFR). In this context, RFD is defined as sustained eGFR of less than 60 mL/min per 1.73 m^2.

Results: During a mean observation period of 5.1 years, 350 participants developed RFD. Significantly, elevated interquartile range (IQR) levels of air pollutants—including PM_{10} (particles ≤10 micrometers, HR 1.389, 95 % CI 1.2-1.606), $PM_{2.5}$ (particles ≤2.5 micrometers, HR 1.353, 95 % CI 1.162-1.575), CO (carbon monoxide, HR 1.264, 95 % CI 1.102-1.451), and NO_2 (nitrogen

dioxide, HR 1.179, 95 % CI 1.021–1.361)—were significantly associated with an increased risk of RFD, after factoring in demographic and health variables. Moreover, exposure to PM_{10}, $PM_{2.5}$, and CO was associated with decreased eGFR.

Conclusions: This study demonstrates a substantial link between air pollution exposure and renal function impairment in primary GN, accentuating the significance of environmental determinants in the pathology of immune-mediated kidney diseases.

(2) Validation of an Acute Kidney Injury Prediction Model as a Clinical Decision Support System

Background: Acute kidney injury (AKI) is a critical clinical condition that requires immediate intervention. We developed an artificial intelligence (AI) model called PRIME Solution to predict AKI and evaluated its ability to enhance clinicians' predictions.

Methods: The PRIME Solution was developed using convolutional neural networks with residual blocks on 183,221 inpatient admissions from a tertiary hospital (2013–2017) and externally validated with 4,501 admissions at another tertiary hospital (2020–2021). To assess its application, we conducted a prospective evaluation using retrospectively collected data from 100 patients at the latter hospital, including 15 AKI cases. AKI prediction performance was compared among specialists, physicians, and medical students, both with and without AI assistance.

Results: Without assistance, specialists demonstrated the highest accuracy (0.797), followed by medical students (0.619) and the PRIME Solution (0.568). AI assistance improved overall recall (61.0% to 74.0%) and F1 scores (38.7% to 42.0%), while reducing average review time (73.8 to 65.4 seconds; $p<0.001$). However, the impact varied across expertise levels. Specialists showed the greatest improvement (recall: 32.1% to 64.3%; F1: 36.4% to 48.6%), whereas medical students' performance improved but aligned more closely with the AI model. Additionally, the effect of AI assistance varied by prediction outcome, showing greater improvement in recall for cases predicted as AKI, and better precision, F1 score, and review time reduction (73.4 to 62.1 seconds; $p<0.001$) for cases predicted as non-AKI.

Conclusion: AKI predictions were enhanced by AI assistance, but the improvements varied according to the expertise of the user.

In conclusion, these studies highlight the potential of data-driven approaches in nephrology. The first study establishes a clear link between air pollution and chronic renal function deterioration, emphasizing the importance of considering environmental factors in long-term kidney disease management. The second demonstrates the value of AI in enhancing acute kidney injury prediction, showcasing the potential of advanced analytics in clinical practice for rapid intervention. These findings pave the way for more

targeted interventions and personalized care strategies in nephrology, potentially improving patient outcomes and quality of life across the spectrum of kidney diseases, from acute injuries to chronic conditions.

Keyword : Chronic Kidney Disease, Acute Kidney Injury, Air Pollution, Machine Learning, Time Series Analysis, Glomerulonephritis

Student Number : 2021-33712

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Kidney diseases encompass a diverse spectrum of conditions that significantly impact patient outcomes, ranging from acute kidney injury (AKI) requiring immediate intervention to various chronic conditions requiring long-term management [1,2]. The global burden of kidney diseases is substantial, affecting millions worldwide and presenting significant challenges to healthcare systems, particularly in regions with limited healthcare resources [3]. These conditions can be interrelated in their progression and management, with acute conditions potentially leading to chronic complications, and various underlying pathologies contributing to progressive kidney dysfunction over time [4].

The effective management of kidney diseases requires understanding multiple risk factors and developing robust prediction strategies. Environmental factors, particularly air pollution, have emerged as significant contributors to various chronic diseases [5,6], including kidney diseases, yet their impact on specific conditions such as primary glomerulonephritis (GN) remains understudied [7,8]. Meanwhile, in acute settings, despite numerous studies on AKI prediction and the potential of artificial intelligence approaches [9], the implementation of proactive management strategies remains limited, highlighting the need for more effective clinical decision support systems.

This research presents two complementary approaches to kidney disease prediction and monitoring through time series data analysis. The first study examines the effects of environmental factors, particularly air pollution, on renal function deterioration in patients with primary glomerulonephritis, utilizing longitudinal data to establish temporal relationships between exposure and outcomes. The second study develops and validates an artificial intelligence-based approach for predicting acute kidney injury within a 48-hour window, utilizing temporal clinical data patterns to enhance predictive capabilities.

While these studies address different types of kidney diseases, they share a common methodological foundation in their use of time series data analysis for disease prediction and monitoring. This unified approach to data analysis, applied across both acute and chronic conditions, provides new insights into prediction, monitoring, and prevention strategies in nephrology. Together, these studies contribute to our understanding of both environmental and clinical factors affecting kidney health, potentially improving the quality of patient care through enhanced prediction and monitoring capabilities.

Chapter 2. Air quality and kidney health: Assessing the effects of PM_{10}, $PM_{2.5}$, CO, and NO_2 on renal function in primary glomerulonephritis[1]

2.1. Study Background

Numerous investigations have scrutinized the associations between environmental exposures and health issues. Air pollution, among various environmental factors, has emerged as a significant environmental risk factor exacerbating chronic diseases, including cardiovascular diseases and diabetes [5,6]. Research has disclosed that pollutants, particularly fine particles, aggravate these diseases mainly through inflammatory responses, oxidative stress, and DNA damage [11]. Furthermore, substantial evidence delineates how air pollutants contribute to the risk of kidney diseases, such as chronic kidney disease (CKD) and End-Stage Renal Disease (ESRD) [7,8].

---

[1] This chapter is based on previously published paper, [10] Yi J, Kim SH, Lee H, et al. Air quality and kidney health: Assessing the effects of pm(10), pm(2.5), co, and no(2) on renal function in primary glomerulonephritis. *Ecotoxicol Environ Saf* 2024;281:116593.

While prevailing studies have primarily concentrated on CKD incidence among the general populace and the progression to ESRD among CKD patients, a research void persistently exists concerning various kidney diseases, each characterized by its unique pathology.

Glomerulonephritis (GN), marked by the inflammation of the kidneys' glomeruli, or filtering units, presents as one such disease [12]. GN may emerge as primary, often idiopathically linked to anomalous immune responses, or as secondary, associated with autoimmune diseases, infections, cancers, or metabolic disorders. Research trends reveal an alarmingly rising GN burden, with a 77% increase in incidence, an 81% rise in prevalence, and a doubling of the mortality rate from 1990 to 2019 [13]. GN pathogenesis encompasses abnormal immune reactions inflicting damage on glomerular endothelial and epithelial cells, vascular tissues, and podocytes. Airborne pollutants, once internalized, can trigger diverse immune mechanisms, culminating in macrophage activation and subsequent damage to endothelial and epithelial cells. It is hypothesized that these pathogenetic mechanisms can interact and exacerbate each other, aggravating glomerular injuries [14,15]. Although inquiries into air pollution's effects on autoimmune diseases, including systemic lupus erythematosus (SLE) and lupus nephritis, exist [16,17], scholarly attention towards the impacts of air pollutants on primary GN, which manifests more prevalently than lupus nephritis, remains scant. Noteworthily, the preceding body of

research on air pollutants and primary GN has predominantly focused on assessing GN incidence, with fewer attempts to elucidate the implications of air pollution on renal function among primary GN patients [18,19]. Our study endeavors to bridge this gap by evaluating the repercussions of air pollutant exposure on renal function in individuals with primary GN.

2.2. Methods

2.2.1. Ethical statement

Aligned with the Declaration of Helsinki principles, this study received approvals from the institutional review boards (IRB) of Seoul National University Bundang Hospital (B-2101-658-104) and Seoul National University Hospital (J-2102-080-1197). The IRBs waived the need for acquiring written informed consents from participants, considering the study's retrospective nature and the minimal risk posed to participants.

2.2.2. Study participants

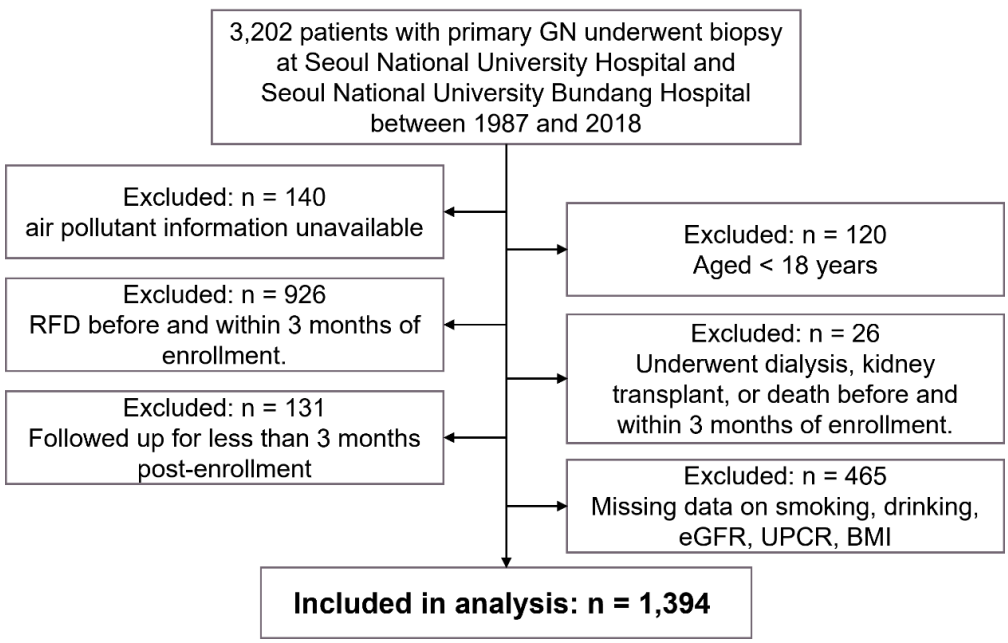
Data for this investigation were derived from two major clinical centers: Seoul National University Bundang Hospital and Seoul National University Hospital, encompassing detailed records of 3,202 individuals diagnosed with primary GN from 1987 to 2018. This cohort included patients with focal segmental glomerulosclerosis (FSGS), immunoglobulin A nephropathy (IgAN), minimal change disease (MCD), membranous nephropathy (MN),

and membranoproliferative glomerulonephritis (MPGN). Primary GN was confirmed through clinical and pathological assessment, and secondary forms of GN were excluded. Diagnostic assessments utilizing light microscopy (LM), immunofluorescence (IF), and electron microscopy (EM) were conducted, with renal pathologists at each institution affirming the diagnoses of primary GN among others. Clinical and demographic data, including age, sex, body mass index (BMI), smoking status, alcohol consumption, pre-existing medical conditions, and laboratory outcomes, were meticulously extracted from hospital information systems, reflecting the information accessible throughout the follow-up period post-biopsy for patient evaluation.

Exclusions were applied to: 140 participants due to the unavailable air pollution data; 120 individuals under 18 years of age; 926 patients experiencing renal function deterioration (RFD) before and within 3 months of enrollment; and 26 subjects who initiated dialysis, underwent a kidney transplantation, or succumbed before and within 3 months post-enrolment. Additionally, 131 patients were excluded due to their follow-up ceasing within 3 months of enrollment, alongside 465 patients lacking data on crucial covariates such as smoking, drinking, estimated Glomerular Filtration Rate (eGFR), urine protein creatinine ratio (UPCR), and BMI. These exclusions aimed to mitigate the influence of early acute fluctuations, facilitating a more accurate longitudinal analysis of air

pollution's chronic impact on renal function. Subsequently, the refined cohort for our analysis constituted 1,394 participants (Figure 1).

Figure 1. Diagram illustrating the selection process of cohort participants.



Primary GN encompasses Focal Segmental Glomerulosclerosis (FSGS), Immunoglobulin A Nephropathy (IgAN), Minimal Change Disease (MCD), Membranous Nephropathy (MN), and Membranoproliferative Glomerulonephritis (MPGN). RFD denotes Renal Function Deterioration.

2.2.3. Air pollution exposure assessment

We utilized air pollution data at a 1 km^2 resolution, in alignment with the methodology of [20], focusing on ground-level concentrations of NO_2, CO, PM_{10}, and $PM_{2.5}$ throughout South Korea from 2002 to 2020. The employed methodology harnessed machine learning algorithms, coupled with Inverse Distance Weighting (IDW) to interpolate air pollution data from the Air Korea database, selecting only those stations maintaining consistent records for a minimum of nine months per year. This approach was enhanced through the integration of satellite data, details concerning land-use such as vegetation indexes and greenness measures, alongside socioeconomic parameters sourced from the Google Earth Engine and the Socioeconomic Data and Applications Center. These multifaceted predictors facilitated a meticulous and high-resolution evaluation of pollution levels [20].

For each participant, we estimated the 1-year average levels of PM_{10}, $PM_{2.5}$, CO, and NO_2 based on their residential addresses, using these pre-calculated monthly air pollution data from the surrounding monitoring stations. In the Cox regression analysis, pollutant levels were updated every 12 months, commencing with the participant's month of entry into the study. In cases where a censoring event preceded the completion of the forthcoming 12-month interval, the average was calculated from the 12 months leading up to the event. Conversely, for the linear mixed model

(LMM) analysis, the air pollutant concentrations for the preceding 12 months were recalculated each time an eGFR measurement was taken, with monthly averages used for the eGFR values. This strategy ensured a tailored exposure assessment for each participant from the point of their inclusion. The monitoring span extended until the earliest of the following occurrences: the onset of RFD, commencement of dialysis or kidney transplantation, death, the conclusion of the 10-year observational period, or the end of our study on December 31, 2020.

2.2.4. Outcome definition

The primary outcome examined in this study was defined as the occurrence of renal function deterioration (RFD), tailored specifically to the clinical context of GN. RFD was determined by a prolonged decrement in the estimated glomerular filtration rate (eGFR) falling below 60 mL/min per 1.73 m^2 for a minimum duration of three consecutive months. Within this timeframe, at least two measurements of eGFR less than 60 were a prerequisite, with none exceeding this threshold, all recorded during outpatient visits. While this criterion is typically associated with chronic kidney disease (CKD), in the context of GN, which inherently involves renal impairment, we employed the designation RFD to precisely describe the progression of renal function decline. The eGFR was calculated utilizing the formula provided by the Chronic Kidney Disease

Epidemiology Collaboration (CKD-EPI):

$$GFR = 141 \times \min(Scr/\kappa, 1)^{\alpha} \times \max(Scr/\kappa, 1)^{-1.209} \times 0.993^{Age} \times 1.018 \text{ [if female]}$$

where Scr represents serum creatinine, κ is 0.7 for females and 0.9 for males, and α is −0.329 for females and −0.411 for males. The 'min' function indicates the lesser value between Scr/κ and 1, while 'max' refers to the greater value between Scr/κ and 1 [21]. A patient meeting and sustaining this eGFR standard for the stipulated period was classified as having experienced RFD.

For linear mixed model analyses, the continuous eGFR values were employed directly as the outcome variable. The objective was to rigorously assess the influence of air pollution on the longitudinal progression of eGFR values.

2.2.5. Statistical analyses

The baseline characteristics of our study participants, encompassing demographic, clinical, and laboratory parameters, were comprehensively summarized. Continuous variables were presented in terms of their mean values and standard deviations, while categorical variables were expressed as counts and percentages. We evaluated the statistical significance of differences using t-tests and chi-squared tests, employing R statistical software (version 4.1.3) for all analyses.

To explore the relationship between the incidence of RFD and

exposure to various air pollutants (PM_{10}, $PM_{2.5}$, CO, NO_2), time-varying Cox regression models were utilized. The focal exposure variable was conceptualized as the 1-year average concentration of each pollutant, treated as a time-varying variable within the study's framework. We examined three distinct models: Model 1, which acted as a crude model, focusing solely on each air pollutant individually; Model 2 incorporated demographic variables such as participants' sex and age at the start of the study; and Model 3 extensively adjusted for additional factors including pre-existing health conditions like diabetes, hypertension, and cancer, alongside baseline laboratory measurements such as eGFR and UPCR. This comprehensive model also considered lifestyle factors—alcohol consumption and smoking status (current or past) and other variables including GN types (FSGS, IGAN, MCD, MN, MPGN), BMI, and the administration of immunosuppressive agents (ISA) or angiotensin receptor blockers (ARB) within one year following enrollment. Seasonality was incorporated as a categorical variable, establishing spring (March to May) as the reference period, and categorizing the remaining seasons accordingly to account for seasonal fluctuations in air pollutant concentrations. We examined the proportional hazards assumption for control variables through Schoenfeld residuals analysis, log-transforming baseline eGFR values to align with this assumption, and stratifying the elaborately adjusted Model 3 by ISA usage.

To discern the individual and combined effects of air pollutants, we implemented two-pollutant models, facilitating an assessment of each pollutant's impact on RFD risk when adjusted for the presence of another pollutant. Subgroup analyses were conducted to identify how certain factors might alter the association between RFD and air pollution. These factors encompassed age, sex, diabetes status, hypertension, cancer, GN types, and BMI categorizations. In subgroups differentiated by age, sex, hypertension, and BMI, adjustments were made for all control variables in the fully adjusted model minus the variables defining the subgroups. For diabetes, cancer, and GN types, adjustments were confined to age and sex, as delineated in Model 2, owing to the limited number of observations. Multiplicative interaction terms were introduced for each factor alongside pollutants in separate models to evaluate potential modification effects, with interactions considered statistically significant at a p-value under 0.05 using the Wald test.

To enhance the robustness of our findings, further analyses were conducted to account for the time-varying nature of critical covariates in the Cox regression model, inclusive of age, drinking status, smoking habits, BMI, and the use of ISA and ARB, thus capturing their evolving impacts on the relationship between air pollution and RFD. Diseases such as diabetes, hypertension, and cancer were treated as static variables to maintain consistency in diagnostic criteria and data availability throughout the study.

Additionally, systolic blood pressure (SBP) was incorporated as a time-varying covariate to thoroughly evaluate blood pressure fluctuations over the study period. Sensitivity analyses employing extended exposure durations of 2 and 3 years for each pollutant were conducted to validate our findings' consistency across longer exposure windows.

The linear mixed model (LMM) approach was employed to investigate the trajectory of outpatient-measured eGFR values in relation to air pollutant concentrations (PM_{10}, $PM_{2.5}$, CO, NO_2). By treating individuals as random effects, we accommodated within-subject variability, allowing for the modeling of both random intercepts and slopes to acknowledge variations in individuals' baseline eGFR and change in eGFR values over time. The primary exposure variable, represented by the 1-year average concentration of each air pollutant, was dynamically computed based on data preceding the month of eGFR measurement. The LMM's data set consisted of these monthly averages, derived from patients' hospital visit records.

Three modeling approaches were pursued. Model 1 integrated time and pollutant averages as fixed effects, with individuals as random effects; Model 2 extended this by including age and sex; and Model 3 embraced seasonality alongside various diseases and lifestyle variables identified at enrollment, mirroring the structure of Model 3 in the time-varying Cox regression framework. To

ensure the modeling process's numerical stability, continuous variables such as age, time, baseline eGFR, and BMI were standardized by centering around the mean and normalizing by standard deviation prior to their incorporation into the models.

2.3. Results

2.3.1. Participant baseline characteristics and average air pollution exposure during follow-up

Table 1 presents the initial characteristics and mean yearly exposure to air pollutants across the entire cohort, segmented according to the incidence of RFD. The study encompassed 1,394 participants, with 350 experiencing RFD throughout an average follow-up duration of 5.1 years. The interval between eGFR assessments averaged 4.1 months (SD 5.21), with a median of 3.04 months. Individuals who developed RFD were, on average, older at baseline compared to those who did not develop RFD (49.3 *vs.* 40.4 years, $p < 0.001$), with no significant disparity in sex distribution between the two categories. Prevalence rates of diabetes and hypertension were considerably higher in the RFD group, accompanied by lower baseline eGFR measures (75.5 *vs.* 97.1 mL/min per 1.73 m^2, $p < 0.001$). Among GN types, IgAN was the most common, representing 56% of the overall cohort. An elevated proportion of subjects with RFD had FSGS and MPGN in comparison

to those without RFD. The follow-up phase marked consistently higher exposures to PM_{10}, $PM_{2.5}$, CO, and NO_2 among the RFD group as opposed to their counterparts, with all differences showing statistical significance (all p-values < 0.001).

Table 1. Baseline characteristics and average of air pollution exposure during the follow-up period for the entire cohort and individuals with/without incident RFD.

Characteristics	Total	Without incident RFD	With incident RFD	*p*-value*
Subjects, n	1,394	1,044	350	
Age at enrollment, mean (SD)	42.6 (15.3)	40.4 (14.7)	49.3 (15.0)	< 0.001
Male, n (%)	715 (51%)	533 (51%)	182 (52%)	0.759
Diabetes, n (%)	131 (9%)	79 (8%)	52 (15%)	< 0.001
Hypertension, n (%)	690 (49%)	467 (45%)	223 (64%)	< 0.001
Cancer, n (%)	48 (3%)	29 (3%)	19 (5%)	0.019
eGFR at enrollment, mean (SD)	91.7 (22.4)	97.1 (21.2)	75.5 (17.8)	< 0.001
UPCR at enrollment, mean (SD)	3287.4 (4882.0)	3216.2 (5071.3)	3499.8 (4267.6)	0.347
Type of GN, n (%)				< 0.001
FSGS	120 (9%)	75 (7%)	45 (13%)	
IGAN	774 (56%)	585 (56%)	189 (54%)	
MCD	177 (13%)	149 (14%)	28 (8%)	
MN	273 (20%)	205 (20%)	68 (19%)	
MPGN	50 (4%)	30 (3%)	20 (6%)	

Characteristics	Total	Without incident RFD	With incident RFD	p-value*
Drinking, n (%)	473 (34%)	371 (36%)	102 (29%)	0.029
Smoking, n (%)				0.009
Current smoker	177 (13%)	116 (11%)	61 (17%)	
Former smoker	152 (11%)	117 (11%)	35 (10%)	
Body mass index (BMI), mean (SD)	24.1 (3.6)	23.9 (3.5)	24.8 (4.0)	< 0.001
Immunosuppressant usage within 1 year from enrollment, n (%)	490 (35%)	355 (34%)	135 (39%)	0.121
ARB usage within 1 year from enrollment, n (%)	1033 (74%)	731 (70%)	302 (86%)	< 0.001
Concentration of PM_{10}, $\mu g/m^3$, mean (SD)	48.0 (7.3)	47.5 (7.1)	50.8 (7.4)	< 0.001
Concentration of $PM_{2.5}$, $\mu g/m^3$, mean (SD)	25.9 (3.4)	25.7 (3.3)	27.0 (3.3)	< 0.001
Concentration of CO, μg/kg, mean (SD)	587.4 (35.7)	585.7 (35.2)	598.0 (36.9)	< 0.001
Concentration of NO_2, μg/kg, mean (SD)	25.4 (4.8)	25.3 (4.9)	26.4 (4.3)	< 0.001

Abbreviations: RFD, renal function deterioration; SD, standard deviation; eGFR, estimated Glomerular Filtration Rate; UPCR, urine protein creatinine ratio; PM_{10}, particulate matter $\leq$ 10 μm in diameter; $PM_{2.5}$, particulate matter $\leq$ 2.5 μm in diameter; CO, carbon monoxide; NO_2, nitrogen dioxide; FSGS, Focal segmental glomerulosclerosis; IgAN, Immunoglobulin A Nephropathy; MCD, Minimal Change Disease; MN, Membranous Nephropathy, MPGN: Membranoproliferative glomerulonephritis; ARB, Angiotensin receptor blocker.

*p-values for difference of no RFD and RFD groups were obtained from t-test and Pearson's Chi-squared test.

2.3.2. Air pollutant exposure and the risk of RFD development

Our examination of the influence of air pollutant levels on RFD risk, as illustrated in Table 2, reveals significant findings from Cox regression analyses. The analysis, spanning three models, consistently identified PM_{10}, $PM_{2.5}$, and CO as factors significantly elevating the hazard ratios (HRs) for RFD, with more pronounced HRs observed in the adjusted models compared to the unadjusted Model 1. Specifically, the fully adjusted Model 3 demonstrates that PM_{10} is linked to an HR of 1.389 (95% CI 1.12–1.606), signifying a marked increase in RFD risk per interquartile range (IQR) increment of 9.3 μg/m³. Furthermore, in Model 3, $PM_{2.5}$ and CO present HRs of 1.353 (95% CI 1.162–1.575) and 1.264 (95% CI 1.102–1.451) respectively, indicating significant RFD risks for IQR increments of 4.4 μg/m³ and 45.8 μg/kg. NO_2 demonstrated statistical significance only in Models 2 and 3, with Model 2 reporting an HR of 1.191 (95% CI 1.031–1.374) and Model 3 reporting an HR of 1.174 (95% CI 1.021–1.361) for an IQR increase of 6.3 μg/kg.

Table 2. Hazard ratios (95 % confidence interval) for Renal Function Deterioration (RFD) by yearly exposure to PM_{10}, $PM_{2.5}$, CO, and NO_2, according to various adjustments.

Exposure	HR(CI) per 1 unit[b]	HR(CI) per IQR[c]
Model 1 (unadjusted)		
PM_{10}	1.024(1.009, 1.040)**	1.252(1.090, 1.438)**
$PM_{2.5}$	1.049(1.015, 1.083)**	1.234(1.068, 1.425)**
CO	1.004(1.001, 1.007)*	1.198(1.044, 1.374)*
NO_2	1.015(0.992, 1.038)	1.098(0.953, 1.265)
Model 2 (adjusted for age and sex)		
PM_{10}	1.036(1.021, 1.052)***	1.390(1.210, 1.597)***
$PM_{2.5}$	1.067(1.033, 1.103)***	1.334(1.155, 1.541)***
CO	1.006(1.003, 1.009)***	1.314(1.145, 1.507)***
NO_2	1.028(1.005, 1.052)*	1.191(1.031, 1.374)*
Model 3 (fully adjusted)[a]		
PM_{10}	1.036(1.020, 1.052)***	1.389(1.200, 1.606)***
$PM_{2.5}$	1.071(1.034, 1.108)***	1.353(1.162, 1.575)***
CO	1.005(1.002, 1.008)**	1.264(1.102, 1.451)***
NO_2	1.027(1.003, 1.050)*	1.179(1.021, 1.361)*

Abbreviations: HR, Hazard Ratio; CI, Confidence Interval; IQR, interquartile range; PM_{10}, particulate matter ≤ 10 μm in diameter; $PM_{2.5}$, particulate matter ≤ 2.5 μm in diameter; CO, carbon monoxide; NO_2, nitrogen dioxide.

[a] Adjusted for age, sex, season, diabetes, hypertension, cancer, eGFR(CKD-EPI) at enrollment, UPCR at enrollment, GN type (FSGS, IgAN, MCD, MN, MPGN), drinking, smoking, BMI, usage of ARBs within 1 year from the enrollment, stratified by usage of immunosuppressants within 1 year from the enrollment.

[b] 1 unit: For PM_{10}, $PM_{2.5}$, CO, and NO, the unit is $\mu g/m^3$, $\mu g/m^3$, $\mu g/kg$, and $\mu g/kg$ respectively.

[c] IQR: For PM_{10}, $PM_{2.5}$, CO, and NO_2, the IQR is 9.3 $\mu g/m^3$, 4.4 $\mu g/m^3$, 45.8 $\mu g/kg$, and 6.3 $\mu g/kg$ respectively.

*$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

In analyses that included two pollutants based on adjustments from Model 3, the investigation into the combined effects of air pollution on RFD risk indicated that both PM_{10} and $PM_{2.5}$ continued to exhibit a substantial influence. Specifically, after adjusting for CO, the HRs for PM_{10} and $PM_{2.5}$ per IQR increase remained significantly elevated at 1.312 (95% CI: 1.119–1.538) and 1.287 (95% CI: 1.097–1.51), respectively. Likewise, with adjustments accounting for NO_2, HRs for PM_{10} and $PM_{2.5}$ were 1.359 (95% CI: 1.168–1.581) and 1.366 (95% CI: 1.169–1.597), respectively, reinforcing their significant impact on RFD risk even in the presence of other air pollutants. Conversely, with adjustments for $PM_{2.5}$, CO and NO_2 demonstrated a positive association with RFD risk with HRs of 1.186 (95% CI: 1.032–1.364) and 1.183 (95% CI: 1.025–1.365), respectively. However, upon adjusting for PM_{10}, the links for CO and NO_2 failed to achieve statistical significance, indicating a nuanced interplay between these pollutants and RFD risk, as detailed in Table 3.

.

Table 3. Two-pollutant models showing hazard ratios for the risk of RFD associated with an interquartile change in each pollutant.

Pollutant	Further adjusted for PM_{10}	Further adjusted for $PM_{2.5}$	Further adjusted for CO	Further adjusted for NO_2
PM_{10} (HR, 95% CI)	-	-	1.312(1.119, 1.538)***	1.359(1.168, 1.581)***
$PM_{2.5}$ (HR, 95% CI)	-	-	1.287(1.097, 1.510)**	1.366(1.169, 1.597)***
CO (HR, 95% CI)	1.143(0.986, 1.324)	1.186(1.032, 1.364)*	-	-
NO_2 (HR, 95% CI)	1.092(0.942, 1.264)	1.183(1.025, 1.365)*	-	-

Abbreviations: HR, Hazard Ratio; CI, Confidence Interval; IQR, interquartile range; PM_{10}, particulate matter ≤ 10 μm in diameter; $PM_{2.5}$, particulate matter ≤ 2.5 μm in diameter; CO, carbon monoxide; NO_2, nitrogen dioxide. All models in this table are based on Model 3 adjustments from the previous analysis. Pollutants with a correlation coefficient of 0.7 or above were excluded to avoid multicollinearity. Correlations among pollutants were as follows: PM_{10} and $PM_{2.5}$, r=0.856; PM_{10} and CO, r=0.407; PM_{10} and NO_2, r=0.277; $PM_{2.5}$ and CO, r=0.296; $PM_{2.5}$ and NO_2, r=0.07; CO and NO_2, r=0.711. *$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

The outcomes from models that included time-varying covariates like age, alcohol consumption, smoking status, BMI, intake of ISA and ARB, and SBP, aligned with those derived from Model 3 and are elaborated in Table 4. Furthermore, the sensitivity analyses that extended exposure durations to 2 and 3 years are documented in Table 5.

Table 4. Extended time-varying covariate analysis of RFD due to 1-year exposure to PM_{10}, $PM_{2.5}$, CO, and NO_2

Exposure	HR(CI) per 1 unit[a]	HR(CI) per IQR[b]
PM_{10}	1.039(1.023, 1.055)***	1.427(1.234, 1.649)***
$PM_{2.5}$	1.080(1.043, 1.118)***	1.404(1.206, 1.636)***
CO	1.006(1.003, 1.009)***	1.305(1.135, 1.501)***
NO_2	1.027(1.004, 1.051)*	1.185(1.026, 1.368)*

Abbreviations: HR, Hazard Ratio; CI, Confidence Interval; IQR, interquartile range; PM_{10}, particulate matter ≤ 10 μm in diameter; $PM_{2.5}$, particulate matter ≤ 2.5 μm in diameter; CO, carbon monoxide; NO_2, nitrogen dioxide.

The model builds upon Model 3 by incorporating updates to key covariates such as age, drinking status, smoking status, BMI, and ISA or ARB usage, transitioning them to time-varying formats to capture dynamic changes over time. Additionally, Systolic blood pressure (SBP) has been newly introduced as another time-varying covariate.

[a] 1 unit: For PM_{10}, $PM_{2.5}$, CO, and NO_2, the unit is μg/m³, μg/m³, μg/kg, and μg/kg respectively.

[b] IQR: For PM_{10}, $PM_{2.5}$, CO, and NO_2, the IQR is 9.3 μg/m³, 4.4 μg/m³, 45.8 μg/kg, and 6.3 μg/kg respectively.

*p < 0.05; ** p < 0.01; *** p < 0.001.

Table 5. Extended time-varying cox regression analysis for RFD with 2-year and 3-year exposures to PM_{10}, $PM_{2.5}$, CO, and NO_2

Exposure	2-year exposure		3-year exposure	
	HR(CI) per 1 unit[b]	**HR(CI) per IQR[c]**	**HR(CI) per 1 unit[b]**	**HR(CI) per IQR[c]**
Model 1 (unadjusted)				
PM_{10}	1.029(1.013, 1.045)***	1.300(1.123, 1.505)***	1.017(1.000, 1.035)*	1.174(1.004, 1.374)*
$PM_{2.5}$	1.057(1.021, 1.094)**	1.277(1.094, 1.490)**	1.044(1.006, 1.083)*	1.210(1.028, 1.426)*
CO	1.004(1.001, 1.007)*	1.187(1.033, 1.363)*	1.003(1.000, 1.006)	1.132(0.983, 1.304)
NO_2	1.014(0.991, 1.037)	1.090(0.946, 1.256)	1.010(0.988, 1.034)	1.066(0.925, 1.230)
Model 2 (adjusted for age and sex)				
PM_{10}	1.042(1.025, 1.058)***	1.461(1.262, 1.692)***	1.032(1.014, 1.049)***	1.335(1.139, 1.565)***
$PM_{2.5}$	1.077(1.040, 1.116)***	1.390(1.191, 1.622)***	1.065(1.026, 1.105)***	1.320(1.120, 1.556)***
CO	1.006(1.003, 1.009)***	1.301(1.133, 1.494)***	1.005(1.002, 1.008)**	1.242(1.079, 1.430)**
NO_2	1.026(1.003, 1.050)*	1.178(1.020, 1.360)*	1.022(0.999, 1.046)	1.149(0.995, 1.328)
Model 3 (fully adjusted)[a]				
PM_{10}	1.042(1.025, 1.059)***	1.464(1.255, 1.708)***	1.031(1.013, 1.050)***	1.330(1.126, 1.57)***
$PM_{2.5}$	1.083(1.043, 1.124)***	1.421(1.206, 1.674)***	1.070(1.029, 1.113)***	1.349(1.133, 1.605)***

Exposure	2-year exposure		3-year exposure	
	HR(CI) per 1 unit[b]	**HR(CI) per IQR[c]**	**HR(CI) per 1 unit[b]**	**HR(CI) per IQR[c]**
CO	1.005(1.002, 1.008)**	1.249(1.087, 1.434)**	1.004(1.001, 1.007)*	1.193(1.036, 1.374)*
NO_2	1.025(1.002, 1.049)*	1.168(1.012, 1.349)*	1.022(0.998, 1.045)	1.144(0.989, 1.322)

Abbreviations: HR, Hazard Ratio; CI, Confidence Interval; IQR, interquartile range; PM_{10}, particulate matter ≤ 10 μm in diameter; $PM_{2.5}$, particulate matter ≤ 2.5 μm in diameter; CO, carbon monoxide; NO_2, nitrogen dioxide.

While the analysis of 2-year exposure covered all 1,394 participants, identical in size to the cohort analyzed for the 1-year exposure, the analysis of the 3-year exposure included only 1,372 participants owing to the constrained follow-up data available prior to 2002.

[a] Adjusted for age, sex, season, diabetes, hypertension, cancer, eGFR (CKD-EPI) at enrollment, UPCR at enrollment, GN type (FSGS, IgAN, MCD, MN, MPGN), drinking, smoking, BMI, usage of ARBs within 1 year from the enrollment, stratified by usage of immunosuppressants within 1 year from the enrollment.

[b] 1 unit: For PM_{10}, $PM_{2.5}$, CO, and NO_2, the unit is μg/m³, μg/m³, μg/kg, and μg/kg respectively.

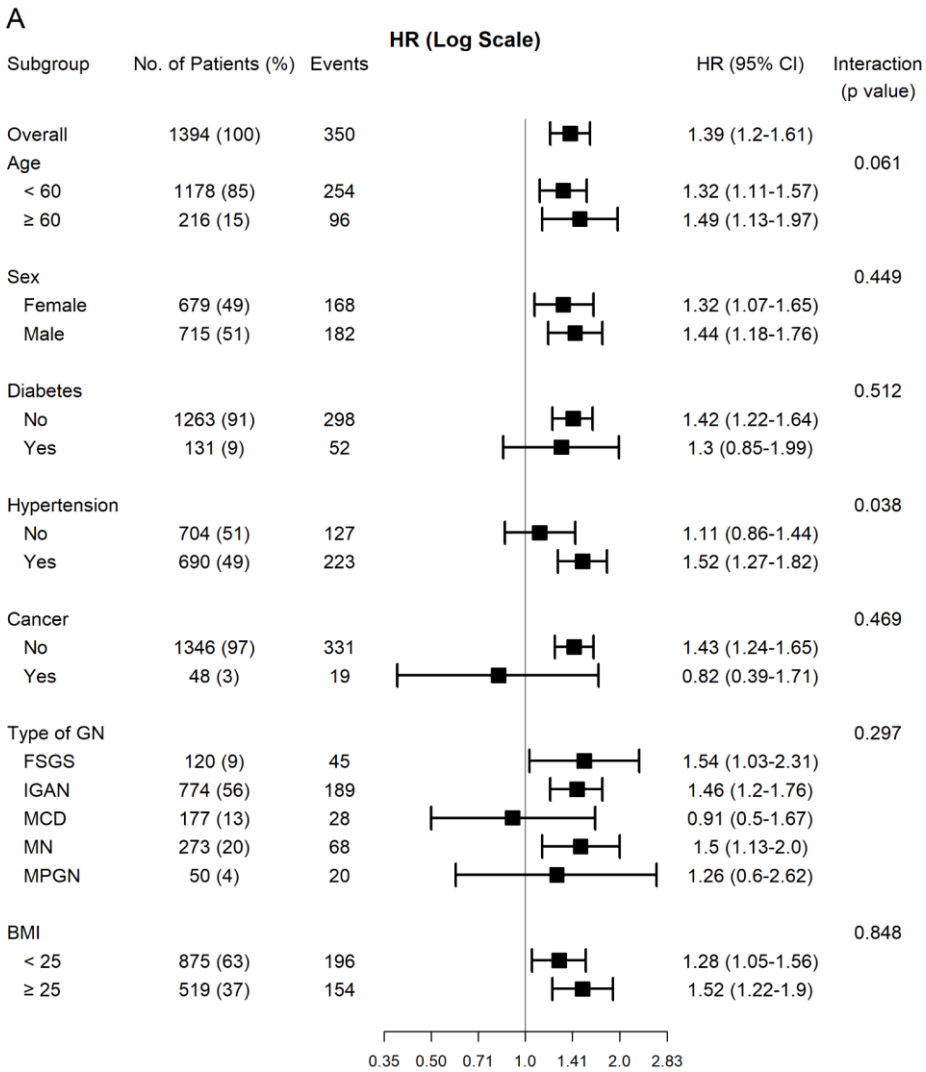
[c] IQR: For PM_{10}, $PM_{2.5}$, CO, and NO_2, the IQR is 9.3 μg/m³, 4.4 μg/m³, 45.8 μg/kg, and 6.3 μg/kg respectively.

*p < 0.05; ** p < 0.01; *** p < 0.001.

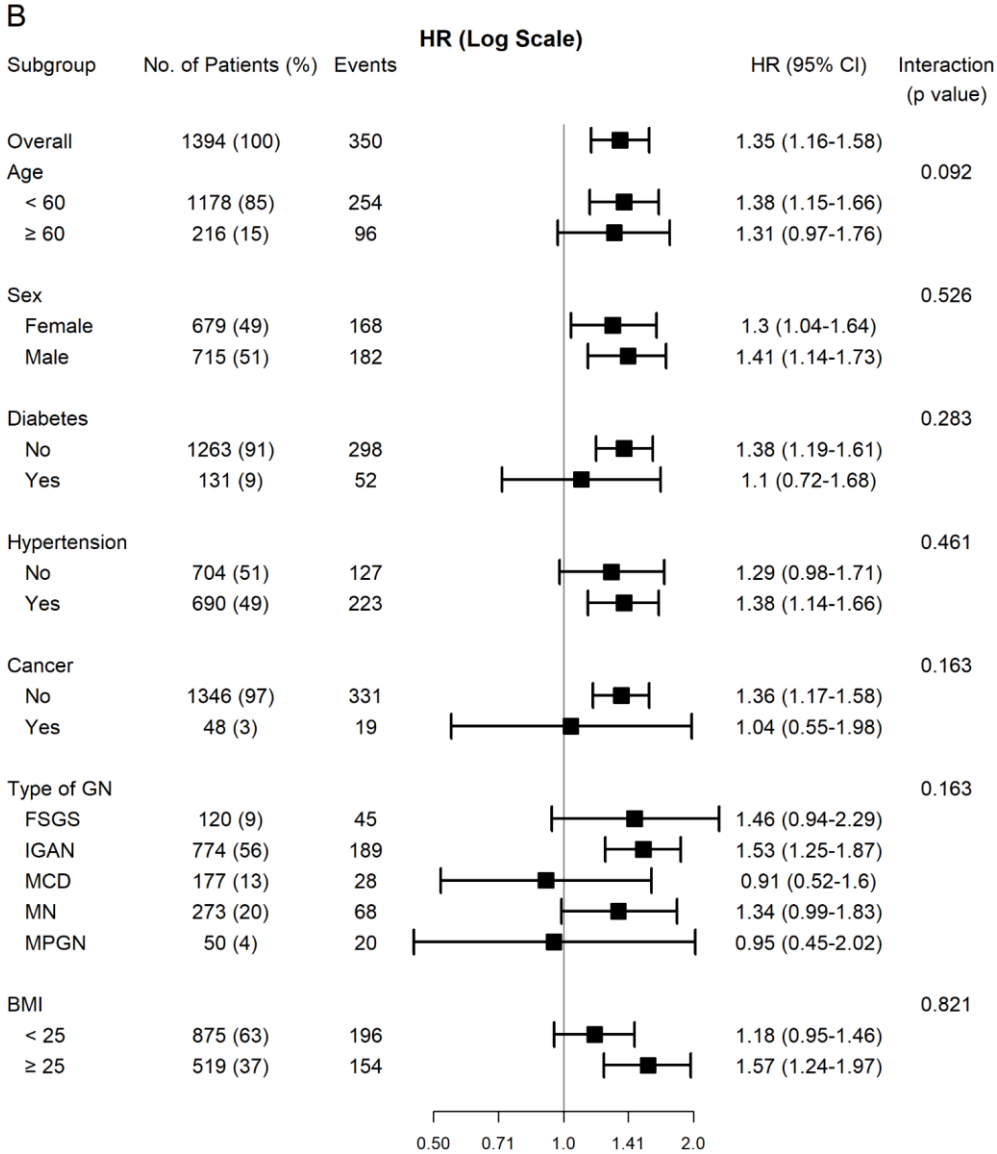
2.3.3. Subgroup analysis of particular matter exposure and risk of RFD

Our subgroup analysis indicated that the increased risk of RFD associated with elevated levels of air pollutants was largely consistent across most subgroups, underscoring an overarching trend of heightened risk. Older individuals (≥ 60 years) exhibited higher HRs for PM_{10} at 1.5 (95% CI:f 1.13–1.99) and for CO at 1.4 (95% CI: 1.05–1.87) compared to younger counterparts, though the interaction p−values did not reveal any significant age−related modification effect on the association between air pollutants and RFD risk. Males demonstrated marginally higher HRs for all pollutants compared to females, with HRs for PM_{10} and $PM_{2.5}$ at 1.44 (95% CI: 1.18–1.76) and 1.41 (95% CI: 1.14–1.73), respectively, albeit without significant interaction p−values. Hypertensive participants showed a distinctly higher HR for PM_{10} at 1.51 (95% CI: 1.26–1.81), with a significant interaction effect (p=0.034), suggesting hypertension could modulate the impact of PM_{10} on RFD. Types of GN did not exhibit consistent patterns across pollutants, and interaction effects remained statistically insignificant. An increased BMI (≥ 25) correlated with higher HRs for all pollutants, evidenced by an HR of 1.5 (95% CI: 1.2–1.87) for PM_{10}, although without significant interaction effects (Figure 2).

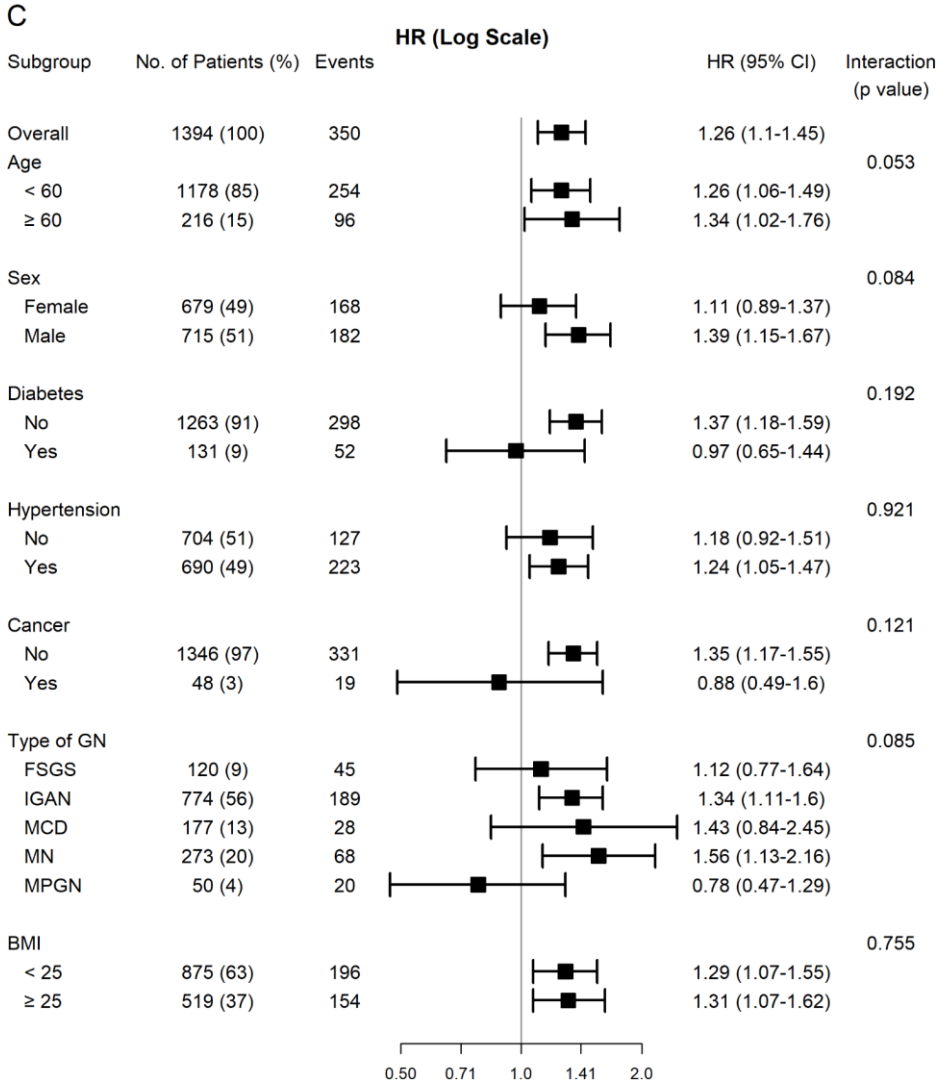
Figure 2. Subgroup analysis of the association between increased particulate matter concentrations and risk of progression to Renal Function Deterioration (RFD).

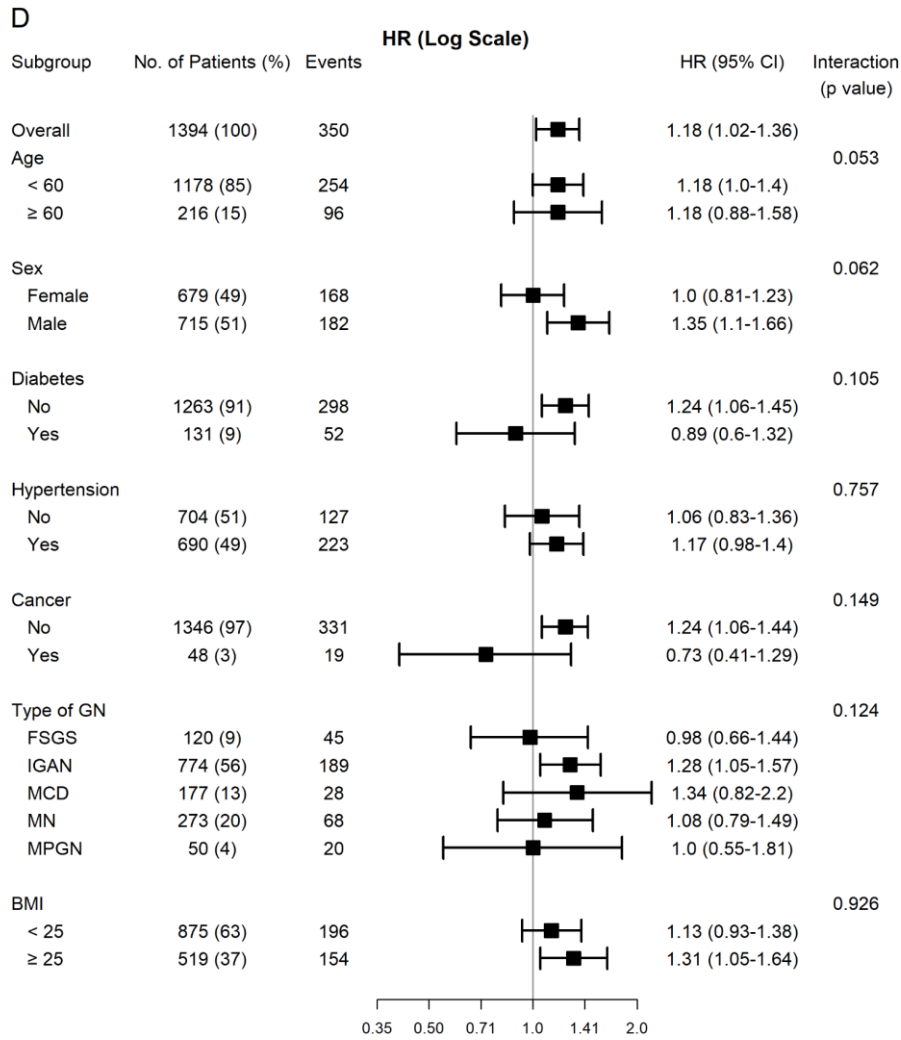


B


HR (Log Scale)
Subgroup
No. of Patients (%)
Events
HR (95% CI)
Interaction (p value)
Overall 1394 (100) 350 1.35 (1.16-1.58)
Age 0.092
< 60 1178 (85) 254 1.38 (1.15-1.66)
≥ 60 216 (15) 96 1.31 (0.97-1.76)
Sex 0.526
Female 679 (49) 168 1.3 (1.04-1.64)
Male 715 (51) 182 1.41 (1.14-1.73)
Diabetes 0.283
No 1263 (91) 298 1.38 (1.19-1.61)
Yes 131 (9) 52 1.1 (0.72-1.68)
Hypertension 0.461
No 704 (51) 127 1.29 (0.98-1.71)
Yes 690 (49) 223 1.38 (1.14-1.66)
Cancer 0.163
No 1346 (97) 331 1.36 (1.17-1.58)
Yes 48 (3) 19 1.04 (0.55-1.98)
Type of GN 0.163
FSGS 120 (9) 45 1.46 (0.94-2.29)
IGAN 774 (56) 189 1.53 (1.25-1.87)
MCD 177 (13) 28 0.91 (0.52-1.6)
MN 273 (20) 68 1.34 (0.99-1.83)
MPGN 50 (4) 20 0.95 (0.45-2.02)
BMI 0.821
< 25 875 (63) 196 1.18 (0.95-1.46)
≥ 25 519 (37) 154 1.57 (1.24-1.97)
0.50
0.71
1.0
1.41
2.0

C


HR (Log Scale)
Subgroup
No. of Patients (%)
Events
HR (95% CI)
Interaction (p value)
Overall 1394 (100) 350 1.26 (1.1-1.45)
Age 0.053
< 60 1178 (85) 254 1.26 (1.06-1.49)
≥ 60 216 (15) 96 1.34 (1.02-1.76)
Sex 0.084
Female 679 (49) 168 1.11 (0.89-1.37)
Male 715 (51) 182 1.39 (1.15-1.67)
Diabetes 0.192
No 1263 (91) 298 1.37 (1.18-1.59)
Yes 131 (9) 52 0.97 (0.65-1.44)
Hypertension 0.921
No 704 (51) 127 1.18 (0.92-1.51)
Yes 690 (49) 223 1.24 (1.05-1.47)
Cancer 0.121
No 1346 (97) 331 1.35 (1.17-1.55)
Yes 48 (3) 19 0.88 (0.49-1.6)
Type of GN 0.085
FSGS 120 (9) 45 1.12 (0.77-1.64)
IGAN 774 (56) 189 1.34 (1.11-1.6)
MCD 177 (13) 28 1.43 (0.84-2.45)
MN 273 (20) 68 1.56 (1.13-2.16)
MPGN 50 (4) 20 0.78 (0.47-1.29)
BMI 0.755
< 25 875 (63) 196 1.29 (1.07-1.55)
≥ 25 519 (37) 154 1.31 (1.07-1.62)
0.50
0.71
1.0
1.41
2.0



Impact of a 1 IQR increase in (A) PM_{10} (particles ≤10 micrometers), (B) $PM_{2.5}$ (particles ≤2.5 micrometers), (C) CO (carbon monoxide), and (D) NO_2 (nitrogen dioxide).

2.3.4. Linear mixed model analysis of eGFR trajectory in relation to air pollutant exposure

We conducted an LMM analysis to evaluate the relationship between eGFR and levels of air pollutants averaged over the 12 months preceding each eGFR assessment. The results revealed that elevations in PM_{10} and $PM_{2.5}$ concentration were significantly associated with reductions in eGFR across models 2 and 3. Specifically, in Model 3, an IQR increase in PM_{10} (9.3 $\mu g/m^3$) corresponded to a yearly decline of −0.653 units in eGFR (95% CI: −1.001, −0.317, $p < 0.001$). A similar pattern was observed for $PM_{2.5}$, though such associations were absent in the unadjusted Model 1. CO showed a consistently negative association with eGFR across all examined models, with Model 3 indicating a significant decline of −2.382 units in eGFR (95% CI: −2.896, −1.868, $p < 0.001$). Conversely, NO_2 did not exhibit a statistically significant relationship with eGFR in any of the models, with Model 3 showing an estimate of −0.533 and a 95% CI ranging from −1.139 to 0.074 (Table 6).

Table 6: Estimated impact on eGFR (95% Confidence Interval) per unit and per IQR increase in yearly exposure to PM_{10}, $PM_{2.5}$, CO, and NO_2

Exposure	Estimate(CI) per 1 unit[b]	Estimate(CI) per IQR[c]
Model 1 (unadjusted)		
PM_{10}	−0.025(−0.065, 0.014)	−0.234(−0.601, 0.132)
$PM_{2.5}$	−0.064(−0.150, 0.022)	−0.282(−0.663, 0.099)
CO	−0.059(−0.073, −0.045)***	−2.714(−3.349, −2.078)***
NO_2	0.044(−0.088, 0.177)	0.278(−0.555, 1.110)
Model 2 (adjusted for age and sex)		
PM_{10}	−0.051(−0.090, −0.012)**	−0.477(−0.839, −0.115)**
$PM_{2.5}$	−0.103(−0.189, −0.018)*	−0.458(−0.834, −0.081)*
CO	−0.068(−0.082, −0.055)***	−3.130(−3.735, −2.525)***
NO_2	−0.080(−0.203, 0.043)	−0.503(−1.277, 0.270)
Model 3 (fully adjusted)[a]		
PM_{10}	−0.070(−0.107, −0.033)***	−0.653(−0.995, −0.311)***
$PM_{2.5}$	−0.119(−0.200, −0.039)**	−0.529(−0.886, −0.171)**
CO	−0.052(−0.063, −0.041)***	−2.382(−2.896, −1.868)***
NO_2	−0.085(−0.181, 0.012)	−0.533(−1.139, 0.074)

Abbreviations: eGFR, estimated Glomerular Filtration Rate; CI, Confidence Interval; IQR, interquartile range; PM_{10}, particulate matter ≤ 10 μm in diameter; $PM_{2.5}$, particulate matter ≤ 2.5 μm in diameter; CO, carbon monoxide; NO_2, nitrogen dioxide.

[a] Adjusted for age, sex, season, diabetes, hypertension, cancer, eGFR (CKD−EPI) at enrollment, UPCR at enrollment, GN type (FSGS, IgAN, MCD, MN, MPGN), drinking, smoking, BMI, usage of ARBs within 1 year from the enrollment, usage of immunosuppressants within 1 year from the enrollment. Time is included as a

random effect in all models.

[b] 1 unit: For PM_{10}, $PM_{2.5}$, CO, and NO_2, the unit is μg/m³ , μg/m³ , μg/kg, and μg/kg respectively.

[c] IQR: For PM_{10}, $PM_{2.5}$, CO, and NO_2, the IQR is 9.3 μg/m³ , 4.4 μg/m³ , 45.8 μg/kg, and 6.3 μg/kg respectively.

*$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

2.4. Discussions

In our study, we meticulously examined the effects of air pollution on renal function by evaluating two primary outcomes: the risk of developing RFD and the decline in eGFR among 1,394 patients diagnosed with primary GN. Our findings reveal significant associations between increased concentrations of PM_{10}, $PM_{2.5}$, CO, and NO_2 and a heightened risk of RFD, as determined by a fully adjusted Cox regression model. Furthermore, our LMM analysis underscored the statistically significant impact of PM_{10}, $PM_{2.5}$, and CO on eGFR decline. Conversely, the effect of NO_2 on eGFR reduction was ambiguous; failing to reach statistical significance in our LMM assessments, albeit one model suggested a subtle trend toward marginal significance.

The significant effects of PM_{10} and $PM_{2.5}$ on renal function, determined through our analysis, align with extant research outlining the harmful repercussions of particulate matter on the incidence of CKD, ESRD, or CKD advancement in the broader populace [7,22]. Pertaining to GN, a recent examination involving 1,979 IgAN patients in China also established a significant

correlation between $PM_{2.5}$ exposure and an escalated risk of kidney failure [23]. While IgAN is recognized as a specific classification of primary GN, our investigation provides a more encompassing understanding of the implications of particulate matter on kidney health within the Korean milieu. Delving into the mechanisms, various studies have elucidated the impact of air pollution on renal function. For example, one investigation showed that inhaled nanoparticles, prevalent in air pollution, can enter the bloodstream, accumulate in vascular lesions, and subsequently undergo filtration and excretion by the kidneys [24]. These pollutants have been identified to accumulate within the kidneys, provoke vascular complications [25], induce oxidative stress and inflammation leading to renal cell apoptosis—the programmed death of kidney cells—and impede autophagy—the cellular mechanism for degrading and recycling cellular components [26]. Exposure to particulate matter has been documented to initiate pathways that result in podocyte injury, a crucial component of the glomerular filtration barrier, accumulating inflammatory cytokines [27]. In individuals afflicted with GN, these adverse effects could manifest more severely due to pre-existing renal vulnerability.

Our observations revealed that while both CO and NO_2 were linked to RFD, the robustness and consistency of these associations varied. CO demonstrated a consistently significant association with RFD across all models. However, the relationship between NO_2

exposure and RFD presented more variability. This association reached statistical significance at the .05 level in Models 2 and 3, yet it was not significant in Model 1. In terms of eGFR decline, CO was consistently associated with reduced eGFR, whereas NO_2 did not exhibit a significant effect. These findings from our study add to the complex narrative found in existing literature concerning the impact of NO_2 on renal health. Research exploring the influence of NO_2 on renal function within the general populace or among patients with non-GN conditions has produced inconsistent outcomes. For instance, a study [28] analyzing data from 26,985 individuals in Poland discovered that increments in medium-term and annual NO_2 and $PM_{2.5}$ levels corresponded with an increased prevalence of CKD patients. Similarly, a study involving 169 older adults identified a linkage between short- and medium-term exposure to NO_2 and a decline in eGFR [29]. Conversely, a study [30] observed that long-term exposure to elevated levels of PM_{10} and SO_2, but not NO_2, associated with diminished eGFR values. A study on Korean adults revealed that long-term exposure to $PM_{2.5}$, PM_{10}, CO, and NO_2 was associated with diminished eGFR levels, albeit effects on CKD were observed only for $PM_{2.5}$ and PM_{10} [31]. Direct evidence linking NO_2 exposure to renal health remains scarce, in contrast to its acknowledged impacts on cardiovascular and pulmonary systems. Nevertheless, the broad biological and pathological pathways of air pollutants still affect kidney health, as noted by [32]. Discrepancies

across studies could arise from different methods of measuring NO_2 exposure (hourly or annually), its combustion-related generation, and the residential proximity of patients to major roadways and industrial facilities, which are significant sources of NO_2, potentially influencing the extent of renal effects observed due to increased exposure levels for those living nearby [33].

Acknowledging the complex mix of various air pollutants in real-world atmospheric conditions, our study utilized two-pollutant models to account for the interaction between different pollutants. Even when controlling for CO and NO_2, significant associations were still evident for PM_{10} and $PM_{2.5}$ with RFD. However, upon adjusting for PM_{10}, both CO and NO_2 portrayed nonsignificant associations. Our results reflect the complexity of pollutant interactions, corroborated by other studies noting shifts in associations for CO and NO_2 when comparing single-pollutant to two-pollutant models. For instance, a study of 8,497 Taipei City residents exploring long-term air pollution exposure effects found that while PM_{10} and $PM_{2.5}$ absorbance remained significantly associated with reduced eGFR in two-pollutant models, NO_2 exhibited only a marginal association [34]. Similarly, in a study with 1,839 participants in Thailand, long-term air pollution impacts showed that the relations between PM_{10} and SO_2 with eGFR persisted in two-pollutant models, whereas CO showed no significant association with eGFR when PM_{10} and SO_2 were included as covariates [30]. Although the

specific mechanisms underlying these patterns remain elusive, studies exploring how air pollution contributes to mortality rates suggest that the effects attributed to NO_2 might surpass its direct harm, hypothesizing that NO_2 acts as an indicator for a wider array of pollutants due to its significant correlation with these pollutants, which are products of combustion. This correlation primarily signals particulate matter as a major contributor to the documented health impacts [35,36].

A subgroup analysis was conducted to identify group-specific differences in the impacts of air pollution on renal function. Increased HRs for RFD were observed associated with PM_{10} and CO exposure in older age groups, suggesting a potential age-related increase in susceptibility. This observation aligns with prior research indicating that air pollution exerts a more substantial impact on the elderly, manifesting as elevated mortality rates and increased signs of oxidative stress and inflammation due to compromised air quality [37-39]. Furthermore, a study conducted in Shanghai highlighted that individuals aged 65 and older were particularly susceptible to the adverse effects of long-term PM_{10} exposure on CKD and the deterioration of eGFR, reinforcing the argument that the senior demographic suffers more significantly from the renal detriments of air pollution [40]. When examining GN types, MCD and MPGN exhibited different responses to various air pollutants. The disparities observed in MPGN could be ascribed to

its limited sample size, potentially diminishing the statistical power and precision. MCD demonstrated lower HRs in response to particulate matter, yet, in contrast, it displayed elevated HRs with CO and NO_2 exposure. Characterized by its ephemeral, steroid-responsive attributes, MCD typically culminates in positive outcomes [41]. A survey involving 580 GN patients revealed that those with MCD experienced the most favorable renal health outcomes and survival rates, likely owing to the condition's benign nature and effective response to corticosteroid treatment [42]. Despite their reversible nature, the higher HRs observed with CO and NO_2 may be attributable to the distinct properties of these pollutants. It is postulated that PM's physical damage and inflammatory response stem from its particle size, whereas CO and NO_2, being gaseous, dissolve into the bloodstream and cause chemical reactions leading to inflammation, thus eliciting divergent responses [43]. Additionally, our study corroborates that a higher BMI is linked to increased HRs for air pollutants, resonating with findings from prior research. A study [44] discovered that long-term exposure to elevated levels of air pollutants significantly exacerbated kidney function in individuals with higher abdominal adiposity. Another investigation underscored that obese adults were particularly prone to the harmful renal outcomes of air pollution, with substances like sulfur dioxide playing a pivotal role, and elevated levels of trans fatty acids in the bloodstream intensifying

the adverse effects [45]. These insights highlight the heightened vulnerability of individuals with obesity to the renal damage induced by air pollution.

This study is not without its limitations. Primarily, it is retrospective in nature, although the prospective enrollment of patients stands as a strength. Secondly, the accuracy of air pollution exposure measurement is subject to potential errors. The estimated ground-level concentrations of pollutants (PM_{10}, $PM_{2.5}$, CO, NO_2) were derived from a highly detailed 1 km × 1 km grid, as delineated by [20]. While this method, which integrates multiple predictors such as satellite data, land use, and socioeconomic variables, provides a high-resolution representation of air pollution, these estimates are based on complex modeling and aggregation processes from multiple data sources. Consequently, despite the granular approach, there may exist some inherent variability in estimation accuracy and methodology. Occasionally, reliance was solely on the last known addresses of the patients, which may have introduced discrepancies in exposure assessments. Moreover, our exposure assessment based solely on residential addresses may not fully capture individual exposure patterns, as people often spend a significant amount of time in different locations, such as workplaces or other places of activity, which may have different levels of air pollution. Additionally, this study's reliance on monthly average values for air pollutants limits the capability to pinpoint short-term

fluctuations that might profoundly affect renal function. Furthermore, certain regional factors such as population density or the ratio of physicians per thousand inhabitants, previously proven to influence kidney function in other investigations [30], were not included in our analysis. This decision was partly due to our cohort's composition, dominated by residents (86%, or 1,200 out of 1,394 patients) from the metropolitan areas of Seoul, Gyeonggi-do, and Incheon. Given this concentration, incorporating more regional factors might not have markedly shifted our outcomes yet could curb their applicability beyond these urban areas. Moreover, the study did not account for lifestyle factors such as diet, physical activity, or dyslipidemia, known to significantly impact kidney function. Variability in diagnostic practices among pathologists across different institutions could have also introduced inconsistencies in participant recruitment and diagnosis accuracy. Acknowledging these constraints, we suggest that future research should explore these dynamics more thoroughly and consider standardizing diagnostic processes, perhaps through the incorporation of secondary reviews, to bolster the robustness and comparability of research outcomes.

In spite of these limitations, our investigation stands out as the first to examine the long-term effects of air pollution on patients with primary GN. By focusing on this distinct cohort, we provide new insights into how different pollutants affect renal function over

time. These findings advance our understanding of the environmental determinants of kidney health and underline the importance of considering air quality in both the management and prevention of kidney disease, providing valuable insights for health care providers and policy makers.

2.5. Conclusions

Our study stands out as the first investigation to examine the long-term effects of air pollution exposure on renal function in patients with primary GN. We observed significant associations between elevated levels of PM_{10}, $PM_{2.5}$, CO, and NO_2 with the progression of RFD, as well as between PM_{10}, $PM_{2.5}$, and CO with lower eGFR, after adjusting for control variables. These findings enhance our comprehension of the environmental factors influencing kidney health and emphasize the necessity of factoring air quality into kidney disease management and prevention strategies, offering critical implications for healthcare professionals and policymakers.

Chapter 3. Validation of an Acute Kidney Injury Prediction Model as a Clinical Decision Support System[2]

3.1. Study Background

Acute kidney injury (AKI) is a prevalent and serious clinical condition that requires immediate management [4]. AKI commonly occurs in hospitalized patients, with a prevalence ranging of 6–18%, and its incidence tends to increase gradually over time during hospitalization [46,47]. Despite numerous studies on early AKI detection, proactive management strategies remain uncommon [46]. The integration of artificial intelligence (AI) into clinical decision support (CDS) systems has emerged as a promising approach to address this gap [47–49]. AI has the potential to enhance predictive performance by identifying patients at higher risk of developing diseases and clinical deterioration and who would benefit from specific management strategies [47]. AI-based CDS systems leverage vast clinical data to provide real-time insights and recommendations that can significantly improve risk assessment

[2] This chapter is based on work accepted for publication in *Kidney Research and Clinical Practice*.

accuracy and patient outcomes [50]. Furthermore, AI can help reduce the cognitive load on clinicians by automating routine tasks and highlighting critical information, allowing healthcare professionals to focus on complex decision-making processes [51].

Building upon our previous work where we developed an AI prediction model for AKI [52], we developed the PRIME Solution (PRedIction and Management of acute kidney injury with Explainable AI). This model not only predicts the occurrence of AKI but also uses layer-wise relevance propagation (LRP) [53] among explainable AI methods, to highlight the most critical factors influencing the predictions. This approach allows physicians to understand the rationale behind the predictions and gain insights into necessary corrective actions.

However, AI-based CDS systems can sometimes generate inaccurate predictions or poorly tailored suggestions, potentially leading to clinician distrust and reduced efficacy [47,54]. We hypothesized that an explainable AI model like PRIME Solution could improve predictive performance and clinician acceptance by providing transparency into its decision-making process. To test this hypothesis, we designed a study to assess the impact of PRIME Solution on AKI predictions made by healthcare professionals of varying expertise levels.

We conducted this study with two primary objectives: first, to compare the predictive performance of our PRIME Solution with

that of physicians and medical students; and second, to evaluate how AI assistance influences the prediction capabilities of these healthcare professionals. By comparing their performance with and without AI assistance, we aimed to assess the efficacy and value of our AI model in enhancing human clinical judgment in AKI prediction.

3.2. Methods

3.2.1. Study Design

This single-center study was conducted at a tertiary hospital in South Korea, Seoul National University Bundang Hospital (SNUBH), from April 2023 to February 2024 and involved a prospective evaluation using patient data collected retrospectively. The study comprised two main phases, with a preliminary phase of AI model development.

Preliminary Phase: Development of the AI model (PRIME Solution). Convolutional neural networks (CNNs) with residual blocks were designed to predict AKI in hospitalized patients. The model was developed using data from a tertiary hospital, Seoul National University Hospital (SNUH), and externally validated using data from SNUBH.

Main Evaluation Phases:

1. AKI prediction without AI assistance (SET1). Clinical evaluators assessed the risk of AKI for 100 patients

selected from the SNUBH dataset without AI assistance.

2. AKI prediction with AI assistance (SET2). Evaluators used the PRIME Solution's predictions, including interpretative analyses of risk factors derived from the model's LRP outputs, to assess the same selected patients as in SET1.

We assessed the impact of PRIME Solution on clinical decisions by comparing evaluations performed with and without AI assistance. The Institutional Review Boards of SNHBH and SNUH approved the model development (IRB no. B-1811-502-004, J-1903-090-1019), external validation (IRB no. B-2205-757-305), and evaluation (IRB no. B-2304-825-304) phases of this study. This study adhered to the 1975 Declaration of Helsinki. The requirement for informed consent from patients was waived because of the retrospective nature of the data and minimal risk posed by the study. Written informed consent was obtained from all participating evaluators.

3.2.2. Data Collection

The development data set comprised 183,221 inpatient admissions from SNUH between 2013 and 2017. The dataset was collected from electronic health records and included data regarding various clinical parameters, laboratory results, and patient demographics. We split these admissions into 70% for training, 15% for validation, and 15% for testing. An additional 4,501 admissions

from SNUBH were used for external validation. Both datasets included adult patients hospitalized for at least 3 days, without prior dialysis, baseline creatinine level less than 4.0 mg/dL, and baseline estimated glomerular filtration rate of 15 mL/min/1.73 m^2 or higher.

3.2.3. Definition of AKI

AKI was defined based on KDIGO (Kidney Disease: Improving Global Outcomes) criteria [55]. Due to the limited availability of urine output data, only serum creatinine criteria were used. AKI was defined as an increase in serum creatinine by ≥0.3 mg/dL or an increase to ≥1.5 times from the baseline value. The baseline creatinine was determined as the minimum value from measurements taken within 14 days before admission. If no measurements were available within this period, we used measurements from within 90 or 180 days before admission. In cases where no pre-admission measurements were available, the minimum value from the admission day was used.

3.2.4. Model Development

We utilized CNNs with residual blocks for AKI prediction, focusing on predicting the onset of AKI. CNNs were chosen because of their proven effectiveness in pattern recognition and classification tasks, particularly in time-series data relevant to AKI prediction [56,57]. Specifically, ResNet was selected owing to its

ability to learn detailed data patterns [58]. The comprehensive details of our model development are presented in the following figures and table: Table 7 provides an overview of the features used in the AKI prediction model with their detailed descriptions; Figure 3 illustrates the systematic data preprocessing workflow implemented for optimal model performance; Figure 4 presents the detailed architecture of our AKI prediction model; and Figure 5 demonstrates the model's performance through ROC curves.

Table 7. Categories and Descriptions of Features Used in the AKI Prediction Model

Category	Features
Demographics[a]	Age, Sex, Body mass index
Clinical Status[b]	ICU admission
Baseline Kidney Function[c]	Baseline creatinine, Baseline eGFR
Comorbidities[d]	Acute myocardial infarction, Congestive heart failure, Peripheral vascular disease, Dementia, Pulmonary disease, Connective tissue disorder, Peptic ulcer, Paraplegia, Renal disease, Cancer, Metastatic cancer, Diabetes mellitus, Diabetes complications, Cardiovascular diseases, Liver disease, Severe liver disease, HIV/AIDS, Hypertension, Acute Kidney Injury
Comorbidity Index[e]	Charlson Comorbidity Index
Medicine Use[f]	Angiotensin-converting enzyme inhibitor, Acyclovir,

Category	Features
	Aminoglycoside, Amphotericin B, Angiotensin receptor blockers (ARBs), Beta blocker, Calcium Channel Blockers, Cisplatin, Colistin, Cyclosporine, Diuretics, Nonsteroidal anti-inflammatory drugs (NSAIDs), Statin, Tacrolimus, Vancomycin, Vasopressor
Laboratory Findings[g]	Albumin, Bilirubin, Blood urea nitrogen, Calcium, Chloride, Creatine kinase, Total CO2, Serum creatinine, C-reactive protein, Glucose, AST, ALT, Hemoglobin, Lipase, Platelet, Potassium, Sodium, Troponin I, White blood cell count
Vital Signs[g]	Systolic blood pressure, Diastolic blood pressure, Heart rate, Body temperature
Surgery Information[h]	General Anesthesia, Non-general anesthesia, Surgery time

[a] Static variables: Age and BMI are continuous; Sex is binary (male = 1).

[b] Static variable: Binary indicating ICU admission at the time of hospital admission.

[c] Static variables: Continuous measurements (creatinine is the minimum value from 6 months prior to admission; eGFR is the maximum value during the same period).

[d] Static variables: All comorbidities are binary, indicating diagnosis before admission based on ICD-10 codes.

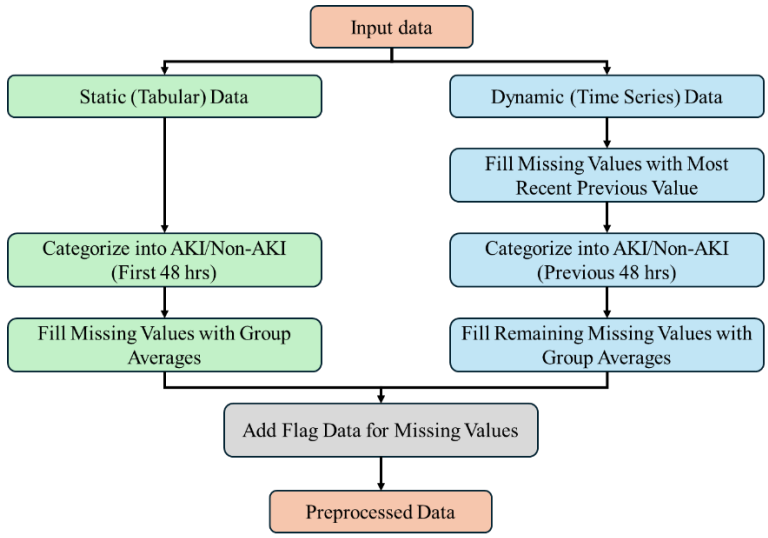
[e] Static variable: Continuous, calculated using the comorbidities.

[f] Static and Dynamic variables: For each medication, a static binary variable indicates prescription within 6 months prior to admission. Dynamic binary variables indicate daily prescription status for the first 7 days of admission.

[g] Dynamic variables: Continuous measurements aggregated into 8-hour intervals for the first 7 days of admission. For Laboratory Findings, the mean value is calculated for each 8-hour period. For Vital Signs, the mean, minimum, and maximum values are calculated for each 8-hour period.

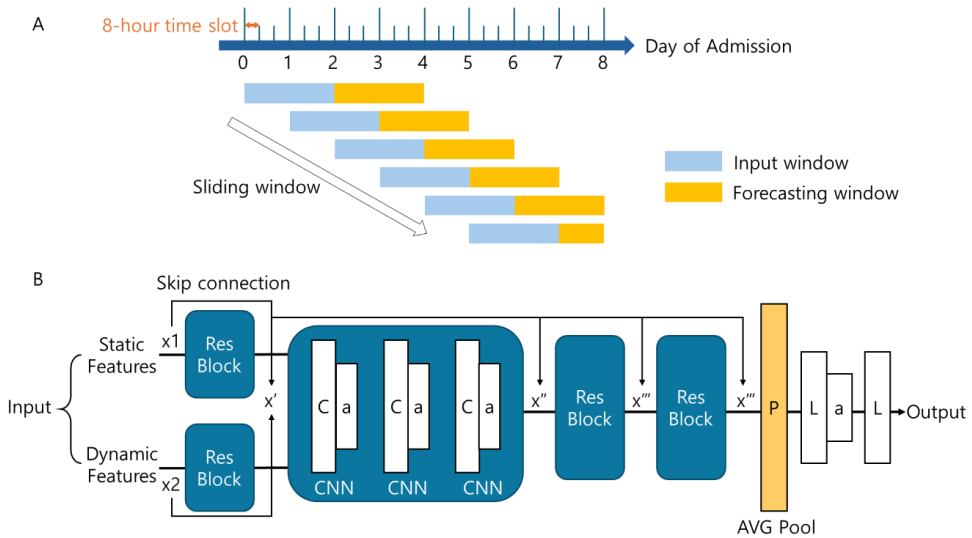
[h] Static variables: General and Non-general anesthesia are binary; Surgery time is continuous.

Figure 3: Data Preprocessing Workflow for AKI Prediction Model.



This flowchart illustrates the preprocessing steps for static (tabular) and dynamic (time series) data used in our AKI prediction model. For dynamic data, missing values are first filled with the most recent previous value when available. Remaining missing values in both static and dynamic data are then handled by categorizing patients into AKI and non-AKI groups (based on the first 48 hours for static data, and the preceding 48-hour window for dynamic data) and filling with respective group averages. Finally, flag data is added to mark initially missing values.

Figure 4: Architecture of the AKI Prediction Model



(A) Sliding Window Approach: The upper part of the figure illustrates the sliding window approach for time series data. It uses a 48-hour input window to predict AKI occurrence within the subsequent 48 hours. The model uses data from the first 7 days of admission to predict AKI occurrence up to the 8th day. Each day is divided into three 8-hour time slots for data aggregation.

(B) Model Architecture: The lower part of the figure depicts the model structure. The model employs dual input processing, separately handling static and dynamic features through initial Residual Blocks (ResBlocks). The core of the model uses a Residual Network (ResNet) structure, with each ResBlock containing three CNN layers. Skip connections are implemented to facilitate gradient flow during training. The output stage consists of an Average Pooling layer followed by Linear layers with activation functions for final prediction.

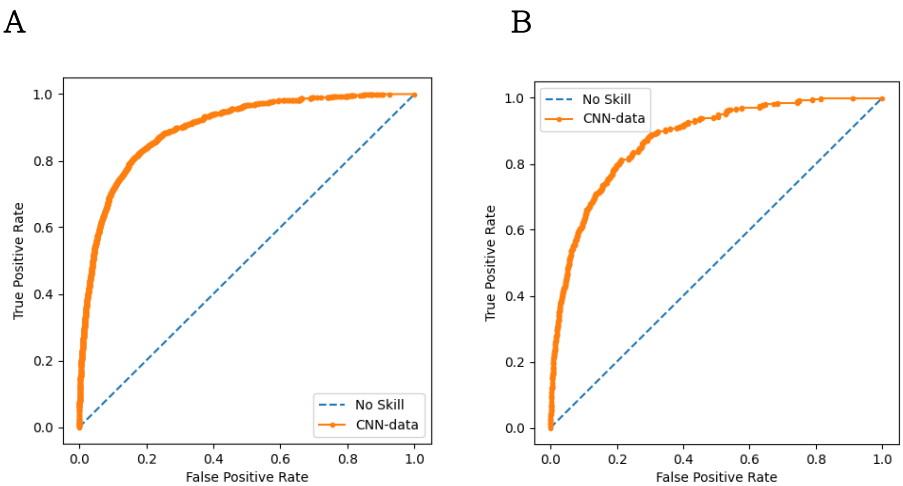
Key components of the model include Residual Blocks (Res Block) for enhancing model depth, Convolutional Neural Network layers (CNN), activation functions (a), an Average Pooling layer (AVG Pool), and Linear layers (L).

The model was trained using a learning rate of 1e-4, a batch size of 256, and the Adam optimizer. Early stopping was implemented using a validation set to prevent overfitting. The total structure comprises five Residual Blocks: two initial blocks for separate input processing and three in the main body as shown in the figure. The hidden dimension was set to 128, and the kernel size for CNN layers was 3.

During model development, we conducted experiments with various hyperparameters, focusing on the ResNet structure's hidden dimension and the number of Residual Blocks. These parameters significantly influence the model's

capacity and performance. We observed that reducing these parameters led to underfitting, while increasing them beyond the current configuration resulted in overfitting and reduced generalization performance. The final configuration (hidden dimension of 128 and five Residual Blocks) was selected as it provided the best balance between model complexity and generalization ability.

Figure 5. Receiver Operating Characteristic (ROC) curves of the PRIME solution model.



(A) ROC curve for the test set (AUC = 0.898). (B) ROC curve for the external validation set (AUC = 0.876). The area under the ROC curve (AUC-ROC) provides a measure of the model's ability to distinguish between AKI and non-AKI cases. An AUC of 1.0 represents a perfect test, while an AUC of 0.5 represents a test no better than random chance.

To determine the optimal threshold for the PRIME Solution, we tested various thresholds for both our original test set and the external validation set. Based on these tests, we set the prediction threshold to 0.9 for this study (Table 8). To enhance model interpretability, we used LRP to determine the contribution of each feature to the predictions of the model (Figure 6).

Table 8. Performance Metrics of the PRIME Solution Model on Test Set and External Validation Set at Selected Thresholds

A

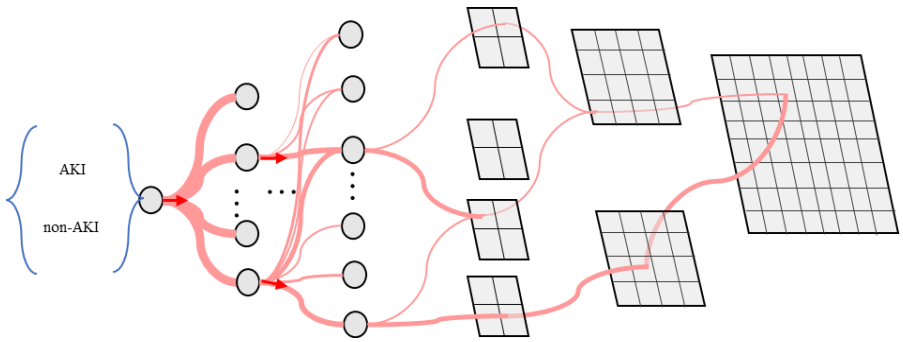
Threshold	Accuracy	Precision	Recall	F1	Specificity
0.3	0.80	0.07	0.84	0.13	0.79
0.5	0.89	0.10	0.73	0.18	0.89
0.7	0.94	0.16	0.56	0.25	0.95
0.9	0.98	0.30	0.24	0.26	0.99

B

Threshold	Accuracy	Precision	Recall	F1	Specificity
0.3	0.53	0.05	0.94	0.09	0.52
0.5	0.70	0.07	0.89	0.13	0.70
0.7	0.83	0.11	0.74	0.18	0.83
0.9	0.94	0.21	0.42	0.28	0.96

(A) Model performance on test set. (B), Model performance on external validation set. The table presents model performance at various thresholds (0.3, 0.5, 0.7, 0.9). The model was initially optimized for high recall to capture a wide range of potential AKI cases. Lower thresholds (e.g., 0.3) demonstrate this high-recall characteristic. For clinical implementation, a threshold of 0.9 was selected to balance the model's sensitive detection capabilities with the need for higher specificity in practical settings. This choice aims to minimize false positives while maintaining the model's ability to identify high-risk cases.

Figure 6: Schematic representation of the Layer-wise Relevance Propagation (LRP) process in AKI prediction model.



The figure demonstrates how relevance is propagated from the outcome (left) to the input features (right). The model distinguishes between AKI and non-AKI cases, with red lines indicating positive relevance and thicker lines representing stronger relevance. LRP is an Explainable AI technique we employed to enhance the transparency of our model. It offers several key advantages: interpretability by mapping model decisions back to input data and highlighting influential features; robustness through the use of Deep Taylor Decomposition to mitigate noise issues common in gradient-based methods; relevance preservation by maintaining relevance values between layers, enabling input-to-output relevance tracing; and improved clinical integration by facilitating trust and integration into clinical workflows through clear explanations. LRP decomposes the model into linear mappings, allowing clinicians to understand which factors most significantly influence the model's predictions, potentially aiding in decision-making and increasing confidence in AI-assisted predictions. By using LRP, we address the 'black-box' nature of deep learning models, making our AKI prediction tool more transparent and interpretable for clinical use. This figure is adapted from Seong et al., "Explainable AKI Prediction with ResNet Toward Real-Time Clinical Decision Support," presented at the KCC XAI 2024 Workshop, part of KCC 2024 (Korea Computer Congress 2024), with permission from the authors.

3.2.5. Data Preparation and Evaluation

For the evaluation phases (SET1 and SET2), we used the data of 100 patients admitted to SNUBH between 2020 and 2021. Twenty patients each were selected from the geriatrics, urology, nephrology, surgery, and orthopedics departments, reflecting the proportion of AKI occurrences across these specialties. The baseline characteristics of these patients are presented in Table 9. Prediction time points were selected as evenly as possible from days 1 to 7 of hospitalization, with an emphasis on the initial days to account for shorter hospitalizations. Only data before AKI onset were included because the goal of the model was to predict the initial occurrence of AKI. Of the 100 patients in the study, 15 developed AKI. During model optimization, 26 patients were excluded due to lack of creatinine data or no change over the course of 48 hours. The remaining 74 patients, including 14 AKI occurrences, were used for most evaluations; however, the data of all 100 patients were used when comparing the performance based on the predictions of the PRIME Solution.

Table 9. Baseline Characteristics and Outcomes of the Model Evaluation Cohort Stratified by AKI Occurrence

Characteristics	AKI (n=15)	No AKI (n=85)	p-value
Age (years), mean (SD)	71.3 (12.1)	67.5 (17.8)	0.434
Male sex, n (%)	8 (53.3)	39 (45.9)	0.801
Body mass index (kg/m2), mean (SD)	25.3 (3.7)	28.2 (44.3)	0.803
ICU admission, n (%)	1 (6.7)	1 (1.2)	0.689
Baseline creatinine (mg/dL), mean (SD)	1.1 (0.6)	1.0 (0.6)	0.767
Baseline eGFR (mL/min/1.73 m2), mean (SD)	79.8 (49.5)	84.0 (51.5)	0.770
Comorbidities			
Acute myocardial infarction, n (%)	0 (0.0)	1 (1.2)	0.325
Congestive heart failure, n (%)	0 (0.0)	3 (3.5)	0.935
Peripheral vascular disease, n (%)	0 (0.0)	3 (3.5)	0.935
Dementia, n (%)	0 (0.0)	6 (7.1)	0.637
Pulmonary disease, n (%)	1 (6.7)	9 (10.6)	1.000
Connective tissue disorder, n (%)	0 (0.0)	1 (1.2)	0.325
Peptic ulcer, n (%)	0 (0.0)	4 (4.7)	0.886
Paraplegia, n (%)	0 (0.0)	0 (0.0)	1.000
Renal disease, n (%)	2 (13.3)	9 (10.6)	0.893
Cancer, n (%)	4 (26.7)	38 (44.7)	0.307
Metastatic cancer, n (%)	0 (0.0)	2 (2.4)	0.689
Diabetes mellitus, n (%)	0 (0.0)	16 (18.8)	0.147

Characteristics	AKI (n=15)	No AKI (n=85)	p-value
Diabetes complications, n (%)	2 (13.3)	4 (4.7)	0.479
Cardiovascular diseases, n (%)	2 (13.3)	19 (22.4)	0.655
Liver disease, n (%)	0 (0.0)	1 (1.2)	0.325
Severe liver disease, n (%)	0 (0.0)	0 (0.0)	1.000
Hypertension, n (%)	0 (0.0)	20 (23.5)	0.080
Acute Kidney Injury, n (%)	4 (26.7)	6 (7.1)	0.062
Charlson Comorbidity Index, median (IQR)	2.0 (0.0-2.0)	2.0 (0.0-3.0)	0.118
Medicine use within six months prior to admission			
Angiotensin-converting enzyme inhibitor, n (%)	0 (0.0)	0 (0.0)	1.000
Acyclovir, n (%)	0 (0.0)	0 (0.0)	1.000
Aminoglycoside, n (%)	0 (0.0)	0 (0.0)	1.000
Amphotericin B, n (%)	0 (0.0)	0 (0.0)	1.000
Angiotensin receptor blockers (ARBs), n (%)	3 (20.0)	10 (11.8)	0.647
Beta blocker, n (%)	1 (6.7)	6 (7.1)	0.621
Calcium Channel Blockers, n (%)	4 (26.7)	17 (20.0)	0.810
Cisplatin, n (%)	0 (0.0)	0 (0.0)	1.000
Colistin, n (%)	0 (0.0)	0 (0.0)	1.000
Cyclosporine, n (%)	0 (0.0)	0 (0.0)	1.000
Diuretics, n (%)	4 (26.7)	12 (14.1)	0.401
NSAIDs, n (%)	9 (60.0)	42 (49.4)	0.634

Characteristics	AKI (n=15)	No AKI (n=85)	p-value
Statin, n (%)	4 (26.7)	8 (9.4)	0.143
Tacrolimus, n (%)	0 (0.0)	0 (0.0)	1.000
Vancomycin, n (%)	0 (0.0)	0 (0.0)	1.000
Vasopressor, n (%)	4 (26.7)	10 (11.8)	0.258
Outcomes	**AKI**	**No AKI**	**p-value**
Incident AKI	15 (100)	0 (0.0)	<0.001
Incident Critical AKI	2 (13)	0 (0.0)	0.016

Binary variables are presented as n (%) and compared using chi-square test. Continuous variables are primarily presented as mean (SD) and compared using t-test. For variables not following normal distribution, data are presented as median (IQR) and compared using Mann-Whitney U test. SD: Standard Deviation; IQR: Interquartile Range.

This study involved the following three groups of 11 evaluators: specialists (one board-certified nephrology subspecialist with 18 years of nephrology training and one internal medicine specialist with 2 years of nephrology training); physicians (one internal medicine specialist without nephrology training and three internal medicine trainees); and medical students (five students from different years of medical school).

Two-stage tests were conducted. For SET1, the evaluators independently reviewed the digital format of the patient data to predict AKI occurrence within 48 hours. After a washout period of 1 to 2 weeks, SET2 was conducted similarly, but with PRIME

Solution predictions, including predictions of the AKI occurrence, top 10 rationales derived from static data, and top 10 rationales derived from dynamic data. The evaluators used a custom evaluation platform with dedicated buttons to record the start and end times for each patient assessment. The platform automatically calculated the evaluation duration for each patient. The evaluators predicted AKI occurrence, selected up to 10 influential variables, and chose appropriate interventions for AKI management. These interventions were categorized into the following six main groups: patient assessment; medication review; imaging studies; hemodynamic stability monitoring; additional tests; and nephrology consultation. Each of these six groups had specific sub-actions (Table 10). The aim was to determine whether the predictions of the PRIME Solution influenced the behaviors of the evaluators. Two specialists finalized the key factors influencing each case of AKI. These key factors were used to evaluate how well the model and the evaluators selected the reasons for predicting the AKI occurrence. The match rate was calculated based on the number of variables chosen by the evaluators that matched the finalized key factors: Match rate = (Number of correctly identified key factors) / (Total number of key factors defined by the specialists).

Table 10: Detailed Interventions for AKI Prediction and Management

Intervention	Definition
1. Patient Assessment	
1−1. Additional Vital Signs Measurement	Check Vital Signs more frequently than the standard interval
1−2. Sepsis Check	Order at least one of: Blood culture, lactate, or procalcitonin
1−3. Bladder Distension Promotion and Scan	Order at least one of: Foley insertion, Clean Intermittent Catheterization (CIC), or Residual Urine (RU) check (preferably after urination) or check nursing records
1−4. Fluid Balance Evaluation (Decreased/Normal/Excessive)	Order input/output (I/O) check (every 1, 4, 6, 8, 12 hours, or daily), body weight check, chest posteroanterior (PA) or anteroposterior (AP) X−rays
2. Medication Review	
2−1. Discontinuation of Nephrotoxic Drugs	Discontinue existing nephrotoxic medication within 24 hours
2−2. Adjustment of Medication Dosage According to Renal Function	Change dosage regimen (dose or frequency) based on renal function
2−3. Fluid Prescription	Prescribe or change fluid (normal

Intervention	Definition
	saline, plasma solution, dextrose with bicarbonate solution, including pre-contrast treatment)
3. Imaging Studies	
3-1. Kidney Ultrasound (CT if necessary)	Order ultrasound (US) of the kidneys, US kidney Doppler, genitourinary CT with or without contrast, abdominal-pelvic CT (APCT) with or without contrast; refer to Radiology or Urology departments
4. Hemodynamic Stability Monitoring	
4-1. Review of Antihypertensive Medication Dosage	Change prescription of ARB, ACEi, CCB, diuretics (Thiazide, Loop diuretics, potassium sparing diuretics), Beta blocker, alpha blocker, vasodilator (discontinue or adjust dosage regimen)
4-2. Anemia Check and Correction	Order iron panel, ferritin, peripheral blood smear, or complete blood count
5. Additional Tests	
5-1. Blood Tests	Order at least one of the following: electrolyte panel, renal panel, urine osmolality, calcium (Ca), phosphorus (P), total carbon dioxide (TCO2), blood urea nitrogen (BUN), creatinine

Intervention	Definition
	(Cr), neutrophil gelatinase-associated lipocalin (NGAL), cystatin C, pH
5-2. Hematuria and Proteinuria Check	Order urinalysis (U/A) using stick and microscopy, 10-panel stick test, or emergency U/A
6. Nephrology Consultation	
6-1. Nephrology Consultation	Refer to Nephrology for unclear causes or stage 3 AKI unresponsive to initial treatment, or dialysis-requiring AKI

Abbreviations—ARB: Angiotensin II receptor blocker, ACEi: Angiotensin converting enzyme inhibitor, CCB: Calcium channel blocker.

3.2.6. Statistical Analysis

We investigated the impact of AI assistance on the prediction performance and evaluated the usefulness of PRIME Solution for AKI prediction. Performance was measured using various metrics, including accuracy, precision, recall, F1 score, specificity, negative predictive value, and false positive rate. Detailed definitions and equations for these metrics are provided in Table 11.

Table 11: Definitions and Equations of Performance Metrics

Metric	Definition	Equation
Accuracy	The proportion of correct predictions among the total number of cases examined	(TP + TN) / (TP + TN + FP + FN)
Precision	The proportion of correct positive predictions out of all positive predictions	TP / (TP + FP)
Recall (Sensitivity)	The proportion of actual positive cases that were correctly identified	TP / (TP + FN)
F1 Score	The harmonic mean of precision and recall	2 * (Precision * Recall) / (Precision + Recall)
Specificity	The proportion of actual negative cases that were correctly identified	TN / (TN + FP)

Metric	Definition	Equation
NPV (Negative Predictive Value)	The proportion of correct negative predictions out of all negative predictions	TN / (TN + FN)
FPR (False Positive Rate)	The proportion of actual negative cases that were incorrectly identified as positive	FP / (FP + TN)

Note: TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative

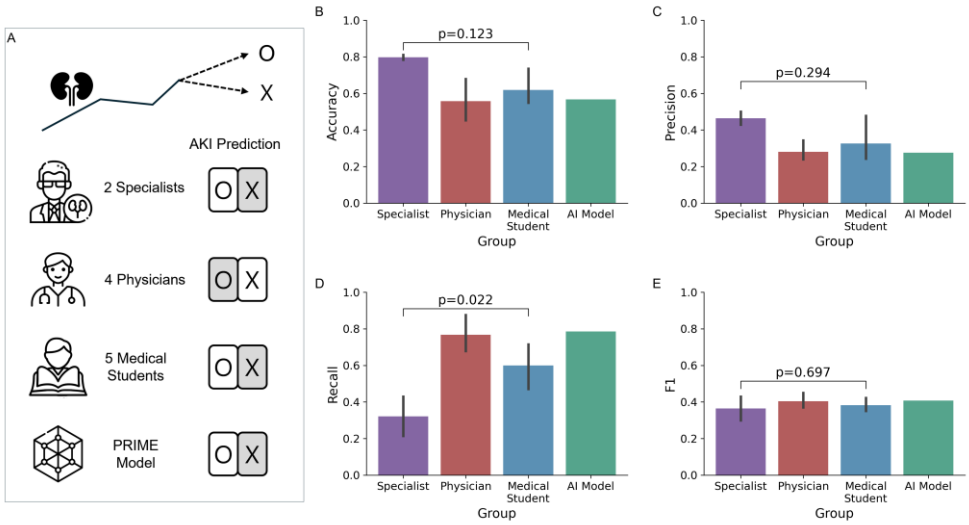
Comparisons within SET1 and SET2 among evaluator groups were performed using one-way analysis of variance (ANOVA). Paired t-tests were performed to compare changes between SET1 and SET2. For analyses of prediction duration and match rates, we utilized all available data points for comparisons between SET1 and SET2 for each evaluator group. Behavioral changes were analyzed by comparing the number of selected actions for each behavior type between SET1 and SET2. All statistical analyses were conducted using R software (version 4.3.2) and Python (version 3.8.16). $P<0.05$ was considered statistically significant.

3.3. Results

3.3.1. Comparison of the Performance Between Evaluators and AI Model (SET1)

The SET1 scenario revealed distinct performance patterns among groups (Figure 7, Table 11, Table 12). Specialists demonstrated the highest accuracy (79.7%) and precision (46.4%) when predicting AKI without AI assistance. The PRIME Solution excelled in recall (78.6%) but had the lowest precision (27.5%), indicating a high rate of both potential AKI case identification and false positives. F1 scores were comparable across all groups (36.4%–40.4%; p=0.697), suggesting a similar balance of precision and recall despite varying individual metrics. Physicians exhibited the second-highest recall (76.8%) after the AI model, while maintaining moderate precision. Medical students performed intermediate, with accuracy (61.9%) falling between that of specialists and physicians.

Figure 7. Comparison of prediction metrics of specialists, physicians, medical students, and the PRIME Solution.



(A) Visual representation of acute kidney injury (AKI) prediction task. (B) Accuracy, (C) precision, (D) recall, and (E) F1 score. Bars represent the average performance of each group. Error bars reflect the 95% confidence intervals. The analysis of variance p-value indicates the statistical significance of differences between groups. Icons are designed by Freepik.

Table 12. Comparison of Prediction Performance Metrics for Acute Kidney Injury of Specialists, Physicians, and Medical Students With and Without AI Assistance

Group	Accuracy	Precision	Recall	F1	Specificity	NPV	FPR
AI model	0.568	0.275	0.786	0.407	0.517	0.912	0.483
SET1 (without AI assistance)							
Specialist	0.797	0.464	0.321	0.364	0.908	0.852	0.092
Physician	0.557	0.281	0.768	0.404	0.508	0.909	0.492
Medical student	0.619	0.326	0.600	0.382	0.623	0.872	0.377
Overall	0.629	0.335	0.610	0.387	0.633	0.882	0.367
ANOVA p-value	0.123	0.294	0.022	0.697	0.082	0.050	0.082
SET2 (AI-assisted)							
Specialist	0.743	0.391	0.643	0.486	0.767	0.902	0.233
Physician	0.510	0.248	0.768	0.371	0.450	0.902	0.550
Medical student	0.586	0.339	0.757	0.434	0.547	0.905	0.453
Overall	0.587	0.315	0.740	0.420	0.552	0.904	0.448
ANOVA p-value	0.201	0.464	0.693	0.412	0.199	0.996	0.199
Comparison of SET1 and SET2							
t test p-value	0.047	0.366	0.045	0.279	0.002	0.286	0.002

Abbreviations: AI, Artificial Intelligence; F1, F1 Score; NPV, Negative Predictive Value; FPR, False Positive Rate; ANOVA, analysis of variance. The presented values represent the mean score of each metric for each group. For detailed explanations of the metrics, see Table 11. The ANOVA p-value represents the results of the ANOVA performed to compare the metrics of the

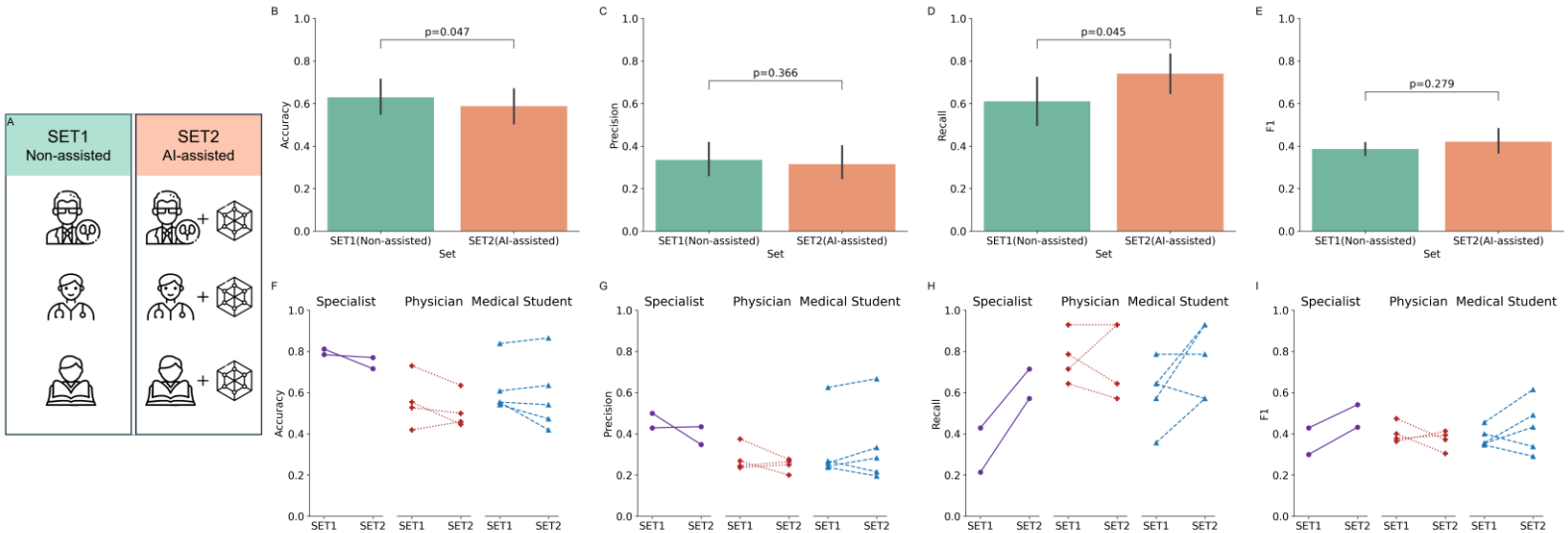
specialist, physician, and medical student groups for each set (SET1 and SET2). The t-test p-value represents the results of a paired t-test performed to compare the metrics for SET1 and SET2.

3.3.2. Comparison of Prediction Performance Metrics According to AI Predictions (SET2)

3.3.2.1. Impact of Assistance Provided by PRIME Solution on Different Skill Levels

AI assistance in SET2 significantly improved overall recall across all groups (from 61.0% to 74.0%; p=0.045) (Table 12, Figure 8), with specialists demonstrating the most substantial increase (from 32.1% to 64.3%). F1 scores also improved (from 38.7% to 42.0%; p=0.279), particularly for specialists (from 36.4% to 48.6%). However, specificity decreased across all groups (from 63.3% to 55.2%; p=0.002), accompanied by increases in false-positive rates (from 36.7% to 44.8%; p=0.002). Figure 8 reveals individual variations within groups: specialists consistently demonstrated improved recall and F1 scores but experienced a slight decrease in accuracy; physicians showed variable precision and recall; and medical students exhibited overall enhancements but did not reach specialist levels of accuracy and precision. These results indicate that AI assistance had varying impacts across different expertise levels, with the most pronounced effects observed in the specialist group.

Figure 8. Comparison of prediction performance metrics with and without the support of the PRIME Solution.



(A) Comparative overview: non-assisted vs. AI-assisted prediction. (B-E) Accuracy, precision, recall, and F1 score of SET1 (non-assisted) and SET2 (AI-assisted). Bars represent the average performance of each set. Error bars reflect the 95% confidence intervals. The p-values indicate statistical significance based on paired t-tests. (F-I) Accuracy, precision, recall, and F1 score of each group with (SET2) and without (SET1) the support of the PRIME Solution. Dots represent the metrics of individual participants, and the lines connect the metrics of the same participant with and without AI assistance. Icons are designed by Freepik.

3.3.2.2. Changes in the Performance According to PRIME Solution Predictions

We examined how the predictions of the PRIME Solution affected the evaluators' performance. When PRIME Solution predicted AKI occurrence, the recall increased significantly, especially among the specialists (from 36.4% to 77.3%). Regarding the predictions of nonoccurrence, the precision and F1 scores improved, notably among specialists (precision: from 20% to 33.3%) and medical students (F1 score: from 21.1% to 40%) (Table 13).

Table 13. Comparison of Prediction Performance Metrics between Specialists, Physicians, and Medical Students according to AI prediction results

Group	Set	Accuracy	Precision	Recall	F1	Specificity	NPV	FPR
AI predicts AKI occurrence (40 Cases, 11 with AKI occurrence)								
Specialist	SET1	0.738	0.533	0.364	0.432	0.879	0.785	0.121
	SET2	0.612	0.395	0.773	0.523	0.552	0.865	0.448
Physician	SET1	0.600	0.391	0.818	0.529	0.517	0.882	0.483
	SET2	0.469	0.322	0.841	0.465	0.328	0.844	0.672
Medical Student	SET1	0.590	0.361	0.636	0.461	0.572	0.806	0.428
	SET2	0.485	0.321	0.782	0.455	0.372	0.818	0.628

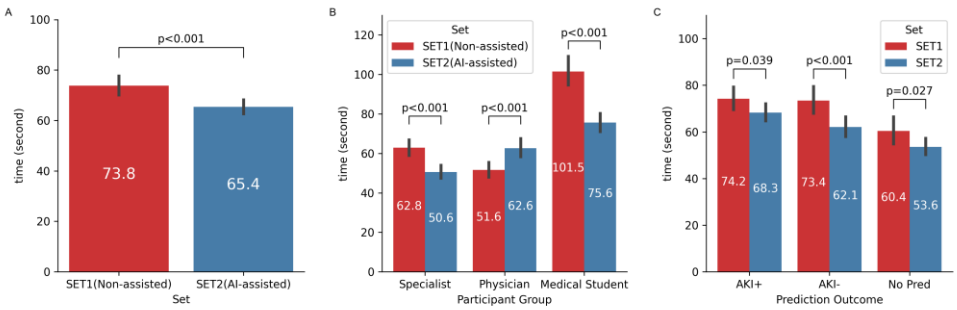
Group	Set	Accuracy	Precision	Recall	F1	Specificity	NPV	FPR
AI predicts NO AKI occurrence (34 Cases, 3 with AKI occurence)								
Specialist	SET1	0.868	0.200	0.167	0.182	0.935	0.921	0.065
	SET2	0.897	0.333	0.167	0.222	0.968	0.923	0.032
Physician	SET1	0.507	0.101	0.583	0.173	0.500	0.925	0.500
	SET2	0.559	0.100	0.5	0.167	0.565	0.921	0.435
Medical Student	SET1	0.653	0.121	0.467	0.192	0.671	0.929	0.329
	SET2	0.706	0.182	0.667	0.286	0.710	0.957	0.290
AI provides no prediction (26 Cases, 1 with AKI occurrence)								
Specialist	SET1	0.981	0.667	1.000	0.800	0.980	1.000	0.020
	SET2	0.846	0.125	0.500	0.200	0.860	0.977	0.140
Physician	SET1	0.644	0.098	1.000	0.178	0.630	1.000	0.370
	SET2	0.596	0.068	0.750	0.125	0.590	0.983	0.410
Medical Student	SET1	0.800	0.138	0.800	0.235	0.800	0.990	0.200
	SET2	0.677	0.089	0.800	0.160	0.672	0.988	0.328

Abbreviations—F1: F1 Score, NPV: Negative Predictive Value, FPR: False Positive Rate.

3.3.2.3. Changes in the Prediction Duration with AI Assistance

AI assistance significantly reduced the decision-making duration from 73.8 seconds to 65.4 seconds ($p<0.001$). This reduction was more pronounced in cases where AI predicted that AKI would not occur (73.4 seconds to 62.1 seconds; $p<0.001$) compared to those where it predicted that AKI would occur (74.2 seconds to 68.3 seconds; $p=0.039$) (Figure 9, Table 14). These results indicated that PRIME Solution improved both the prediction performance and evaluation speed.

Figure 9. Comparative analysis of the duration of the acute kidney injury (AKI) prediction: Evaluating the efficiency of the PRIME Solution's assistance



(A) Comparison of the mean durations of the AKI prediction tasks with and without the aid of PRIME Solution, indicating a difference in time efficiency. Analysis was conducted on 74 out of 100 cases where PRIME Solution provided predictions. (B) Analysis of the mean prediction durations of specialists, physicians, and medical students, showing variations between with and without AI assistance. Analysis was conducted on 74 out of 100 cases where PRIME Solution provided predictions. (C) Outlines of the mean prediction durations when the PRIME Solution predicted the occurrence of AKI (AKI+), predicted no occurrence of AKI (AKI−), and did not offer a prediction (No Pred).

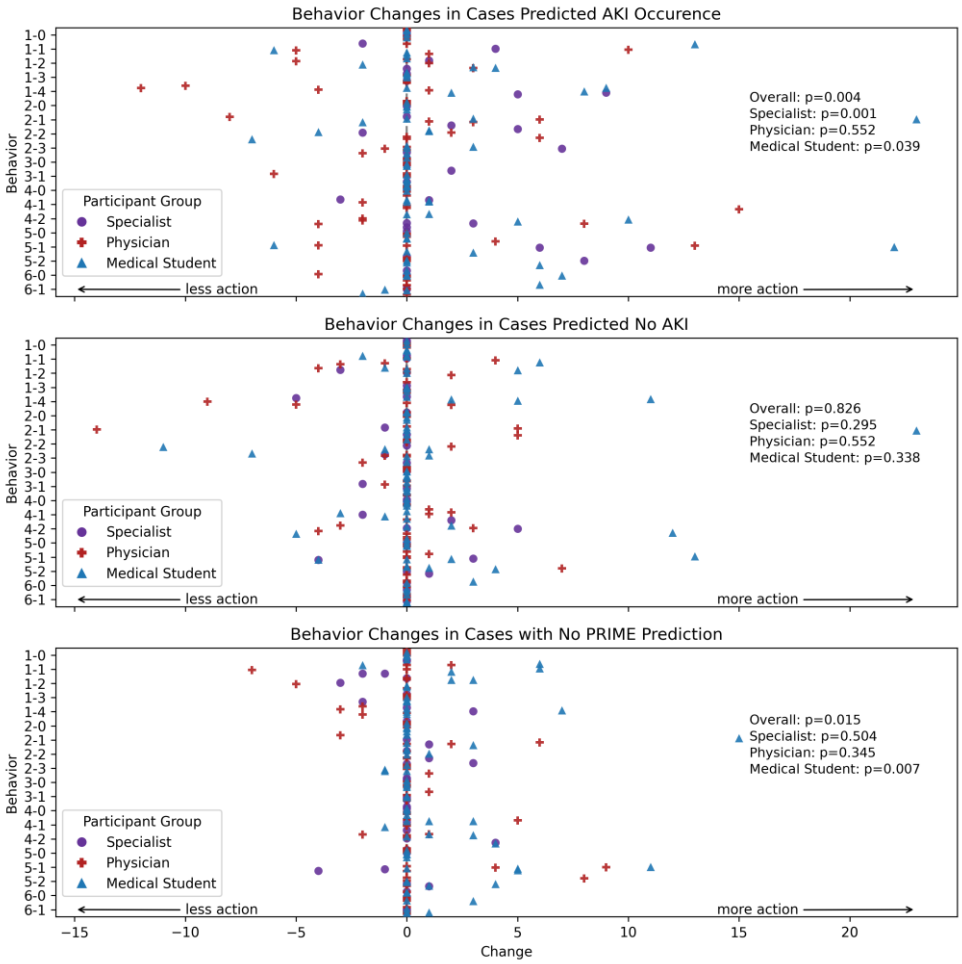
Table 14. Individual and Group Analysis of Average Prediction Times for AKI with AI Assistance

Group	Average Time in SET1 (Non-assisted) (second)	Average Time in SET2 (AI-assisted) (second)	Average Time difference (second)	P-value
Individual				
Specialist S1	64.8	48.0	16.8	0.000
Specialist S2	60.9	53.2	7.7	0.034
Physician P1	59.6	71.4	-11.9	0.088
Physician P2	47.5	41.4	6.1	0.298
Physician P3	39.0	71.2	-32.1	0.000
Physician P4	60.2	66.6	-6.4	0.242
Medical Student M1	86.2	56.9	29.4	0.000
Medical Student M2	116.3	89.0	27.4	0.004
Medical Student M3	131.7	114.7	17.0	0.051
Medical Student M4	71.9	41.9	29.9	0.000
Participant Group				
Specialist	62.8	50.6	12.3	0.000
Physician	51.6	62.6	-11.1	0.000
Medical Student	101.5	75.6	25.9	0.000
Model Prediction				
AKI Occurrence	74.2	68.3	5.9	0.039
No AKI Occurrence	73.4	62.1	11.3	0.000
No Prediction	60.4	53.6	6.9	0.027
Overall				
Overall	73.8	65.4	8.4	0.000

3.3.2.4. Behavioral Changes in Response to AI Predictions

The PRIME Solution influenced the selection of clinical actions across all evaluator groups (Figure 10, Table 15). When the AI predicted AKI occurrence, participants tended to choose more actions, particularly fluid-related evaluations and additional testing. This increase was most notable among specialists and medical students. In contrast, predictions of nonoccurrence resulted in fewer changes in clinical action selections.

Figure 10. Changes in selected clinical actions: Impact of PRIME Solution's assistance



Variations in clinical actions selected by specialists, physicians, and medical students are observed, with each point indicating the change in the number of actions that an individual chose to take. An upward shift (+direction) in points represents an increase in selected actions in clinical decision-making in scenarios with AI assistance compared to those without AI assistance. The figure is divided into three panels, each representing different prediction scenarios by PRIME Solution: (1) cases where AKI occurrence was predicted, (2) cases where no AKI was predicted, and (3) cases where no prediction was provided. This layout allows for comparison of how participants' action selections changed across different prediction contexts.

Table 15: Mean Differences in Clinical Behaviors by AKI Prediction and Clinician Group

Behavior	AKI Predicted			No AKI Predicted			No Prediction		
	Spec	Phys	Med	Spec	Phys	Med	Spec	Phys	Med
1−0	0.00	0.00	−0.25	0.00	0.00	0.00	0.00	−0.25	0.25
1−1	1.00	1.25	1.75	0.00	0.00	−0.25	−1.50	−1.25	3.00
1−2	2.50	−0.25	1.75	−1.50	0.00	0.00	−1.50	−0.25	1.00
1−3	0.00	−0.50	−0.25	0.00	1.00	−0.50	−1.00	0.50	0.00
1−4	7.00	−6.25	4.50	−3.50	−3.00	4.25	1.00	−1.50	2.00
2−0	0.00	0.25	0.00	0.00	0.00	−0.25	0.00	0.00	0.00
2−1	1.00	0.50	6.00	−2.50	−1.00	5.25	0.00	1.75	4.00
2−2	1.50	2.50	−2.25	0.00	−0.25	−2.50	0.50	0.50	−1.00
2−3	10.00	−2.50	0.00	−0.50	−1.25	−1.50	1.50	−0.25	−0.75
3−0	−0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3−1	1.50	1.25	−0.75	−1.00	0.00	−0.50	−0.50	1.25	0.00
4−0	0.00	−0.50	0.00	0.00	0.00	−0.25	0.00	0.00	0.00
4−1	−1.00	3.00	−0.50	0.00	1.00	−1.75	0.00	1.25	−0.50

Behavior	AKI Predicted			No AKI Predicted			No Prediction		
	Spec	Phys	Med	Spec	Phys	Med	Spec	Phys	Med
4−2	3.00	0.00	3.50	3.00	−1.00	1.75	2.00	−1.00	1.25
5−0	0.50	0.00	0.25	0.00	0.00	0.00	−0.50	0.00	0.00
5−1	8.50	3.25	4.75	−0.50	0.00	2.75	−2.50	2.50	5.25
5−2	4.00	6.50	1.50	0.50	1.25	0.75	0.50	2.00	1.25
6−0	0.00	−1.00	1.75	0.00	−0.25	0.75	0.00	−0.50	0.75
6−1	1.00	−0.50	0.25	0.00	0.00	1.00	−1.00	0.00	0.75
P−value	0.001*	0.552	0.039*	0.295	0.552	0.338	0.504	0.345	0.007*

Behavior codes: 1−0: Patient Assessment, 1−1: Additional Vital Signs Measurement, 1−2: Sepsis Check, 1−3: Bladder Distension Promotion and Scan, 1−4: Fluid Balance Evaluation 2−0: Medication Review, 2−1: Discontinuation of Nephrotoxic Drugs, 2−2: Adjustment of Medication Dosage, 2−3: Fluid Prescription 3−0: Imaging Studies, 3−1: Kidney Ultrasound 4−0: Hemodynamic Stability Monitoring, 4−1: Review of Antihypertensive Medication, 4−2: Anemia Check and Correction 5−0: Additional Tests, 5−1: Blood Tests, 5−2: Hematuria and Proteinuria Check 6−0: Nephrology Consultation, 6−1: Nephrology Consultation.

Clinician groups: Spec: Specialist, Phys: Physician, Med: Medical Student.

Values represent mean differences in the number of behaviors between Set 2 and Set 1 (Set 2 − Set 1). Positive values indicate an increase in the behavior, while negative values indicate a decrease.

Clinician groups: Spec: Specialist, Phys: Physician, Med: Medical Student.

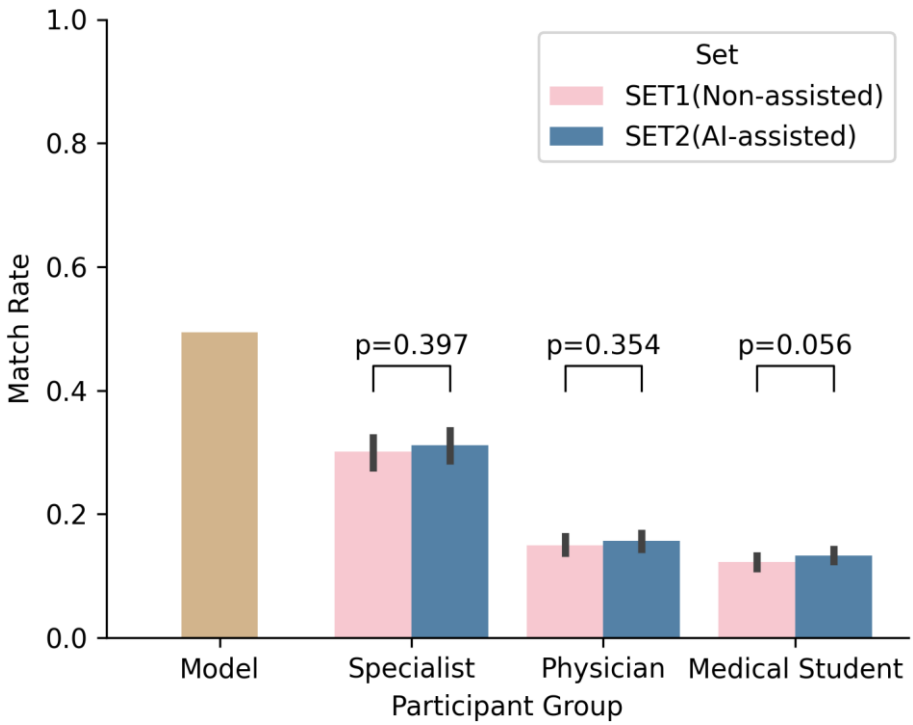
P−values are calculated using paired t−tests for each group and prediction category.

*$p < 0.05$, statistically significant.

3.3.2.5. Model Explanation Match Rate

The match rates of key predictive variables identified by the PRIME Solution through LRP analysis and those selected by the evaluators were compared (Figure 11). The PRIME Solution, which selected 20 variables, showed a higher match rate because of its broader selection set. The evaluators, who were limited to choosing up to 10 variables, had slightly improved match rates with AI assistance (SET2) compared to those without AI assistance (SET1). For specialists, physicians, and medical students, the match rates with AI assistance increased; however, these improvements were not statistically significant ($p=0.397$, $p=0.354$, and $p=0.056$, respectively).

Figure 11. Comparison of match rates of the key predictive variables determined by the evaluator groups and the model.



The match rates of key predictive variables identified by the PRIME Solution and those selected by various evaluator groups when making AKI predictions. The model, which selects 20 variables, naturally shows a higher match rate because of its broader selection set compared to that of the evaluators who were limited to choosing up to 10 variables. SET1 represents the non-assisted selection (SET1), and SET2 indicates AI-assisted selection (SET2).

3.4. Discussion

This study demonstrated the capacity of PRIME Solution to enhance AKI prediction by integrating AI into CDS systems, benefiting clinicians with various expertise levels. Prior to the integration of AI technologies, AKI risk prediction was based on statistical methods and baseline patient data collected before clinical events or interventions. Recent advancements in AI have revolutionized this approach by incorporating not only baseline data but also dynamic real-time data collected during hospitalization, thereby significantly improving predictive capabilities [59]. Unlike previous studies that used traditional statistical models [60,61], our approach leveraged the ResNet architecture, which is specifically tailored to capture temporal patterns in time-series data. This advanced CNN model effectively modeled the evolving nature of patient data collected during hospitalization, thereby improving the performance of AKI risk prediction. However, the inherent complexity of deep learning models, often referred to as the "black box" problem, limits transparency in the decision-making processes [62]. To address this, LRP was utilized and offered interpretable insights regarding the predictions of the model by identifying the most influential variables. This not only mitigated the "black box" issue but also increased user trust and supported more informed clinical decisions.

Although numerous studies have developed AKI prediction

models, evaluations in the context of physician decision-making are scarce. One study directly compared the performance of an AKI prediction model with that of physicians at the time of admission to the intensive care unit and reported areas under the receiver-operating characteristic curve of 0.80 for physicians and 0.75 for the model [63]. Our study further compared predictive metrics for AKI prediction with and without AI support among specialists, physicians, and medical students. This comprehensive investigation illustrated that AI enhanced decision-making in clinical settings, significantly improving predictive performance and efficiency among those with various levels of medical expertise.

The PRIME Solution was designed with a strong emphasis on detecting the onset of AKI, a critical factor for early intervention. This focus resulted in high recall (78.6%), effectively identifying patients who actually developed AKI, but the lower precision (27.5%) indicated a higher occurrence of false alarms. This trade-off demonstrates the ability of the model to detect potential AKI cases while highlighting the need for refinement to reduce unnecessary alerts without compromising sensitivity.

Integrating the PRIME Solution into AKI prediction significantly improved prediction performance, especially recall across all groups. Specialists and medical students demonstrated improvements in F1 scores, indicating a better balance between precision and recall. However, these improvements came with decreased specificity, as

the model prioritized identifying true positives even at the risk of increasing false positives. This approach was used to ensure that potential AKI cases were not missed, enabling early intervention. Interestingly, AI assistance had varying impacts depending on the evaluator's expertise level. Specialists showed a substantial improvement in recall and F1 scores while maintaining higher accuracy than the AI model, indicating that they effectively integrated AI support with their clinical judgment. Medical students also demonstrated substantial improvements; however, their performance was more similar to that of the AI model, suggesting that they relied more heavily on the AI's recommendations without critically evaluating them to the same extent as more experienced clinicians. Physicians exhibited the most individual variability: the performance of some improved with AI assistance, whereas that of others declined. These findings align with research on human–AI collaboration, revealing that the benefits of AI support depend on users' expertise [64]. Experts may process AI explanations with less cognitive load by drawing on their subconscious domain knowledge, potentially allowing them to better assess AI predictions' uncertainty and accuracy. Conversely, those with less expertise might struggle to extract meaningful insights from AI explanations [65].

The PRIME Solution's prediction outcomes significantly influenced predictive performance and clinical behavior. When AI

predicted AKI occurrence, recall markedly increased, especially among specialists (from 36.4% to 77.3%), suggesting that AI predictions can effectively guide clinicians in identifying potential AKI cases, thereby improving early interventions. Conversely, precision and F1 scores improved for non-AKI predictions, highlighting AI's role in supporting more accurate negative diagnoses and its potential use as a screening tool to exclude low-risk patients. Additionally, PRIME Solution led to increased clinical actions when predicting AKI occurrence, particularly among specialists and medical students. This increase in clinical action selections is likely to translate into better patient outcomes, such as higher recovery rates and reduced AKI severity. This proactive approach allows for timely interventions, which are critical when managing AKI. Future studies should focus on quantifying these impacts using specific clinical metrics [66,67]. It is important to note that the selective improvement in performance, varying with both the presence and content of AI predictions, strengthens the argument that PRIME Solution genuinely enhanced clinical decision-making. Although the sample size was limited, the observed decline in predictive performance when AI predictions were not provided suggests that the results of this study were not merely a consequence of learning effects associated with repeated evaluations.

PRIME Solution demonstrated practical benefits in clinical

settings, such as decreased average review duration, with the most significant time savings observed in non-AKI cases (73.4 to 62.1 seconds, $p<0.001$) compared to cases where AKI was predicted (74.2 to 68.3 seconds, $p=0.039$). This efficiency gain aligns with previous studies [68,69] and underscores the practical benefits of AI assistance, allowing clinicians to allocate more time to complex cases and other critical tasks. The reduction in review time with AI assistance is particularly valuable in urgent clinical scenarios in which quick and accurate decision-making is necessary. This efficiency can enhance overall clinical workflow, particularly in time-sensitive situations. By streamlining the prediction process, AI-CDS systems can potentially alleviate clinical load, enabling faster decisions without compromising prediction accuracy [66,70]. Furthermore, the match rates of key predictive variables showed a slight improvement with AI assistance, although the difference was not statistically significant (specialists: $p=0.397$; physicians: $p=0.354$; medical students: $p=0.056$). This suggests that AI can help clinicians select more relevant variables, potentially leading to more informed and accurate assessments. However, the methods of presenting AI insights, such as specific interface designs or visualization techniques, need further refinement to enhance this aspect more effectively.

This study has limitations inherent to its design and methodology. The single-center nature raises questions about the

generalizability of the findings. Differences in patient demographics, clinical practices, and healthcare infrastructures across institutions may lead to variable outcomes when implementing the PRIME Solution elsewhere. Additionally, the study's reliance on a relatively small sample of 100 patients and a limited number of evaluators may not provide a comprehensive overview of the effectiveness across broader clinical settings. Variability in baseline AKI assessment accuracy among evaluators with different expertise levels introduces another layer of complexity that potentially influences the perceived impact of the PRIME Solution. Moreover, the study did not fully explore the dynamic nature of clinical environments, where patient conditions and clinical decision-making factors can evolve rapidly, limiting the applicability to real-world settings. Finally, our AI model's focus on high sensitivity led to decreased specificity, resulting in more false positives. While high recall is critical for early detection and prevention of AKI, this approach could increase clinician workload and potentially lead to alert fatigue in practice, compromising the model's effectiveness.

In conclusion, our study illustrated the promising role of AI in improving the prediction of AKI, particularly through enhancements in recall and efficiency. Nonetheless, integrating AI into clinical practice must be approached with caution, ensuring that such systems augment clinicians' judgment without undermining it. Future research should focus on refining AI models to achieve an

optimal balance between sensitivity and specificity, explore the psychological and behavioral impacts of AI on clinical decision-making, and develop educational strategies to maximize the benefits of AI-CDS systems across all levels of medical expertise. Additionally, future studies should include multicenter trials with larger patient cohorts and diverse hospital types, patient demographics, and clinical practices to validate the model across various healthcare settings. Furthermore, they should aim to develop AI models tailored to specific clinical contexts and consider differences in the prevalence of AKI and its outcomes in various clinical settings, such as medical and surgical intensive care units [71]. The different etiologies of AKI, including sepsis, surgery, and contrast agents should also be evaluated [50,72,73]. Addressing potential barriers to clinical implementation is crucial. A key consideration is the dynamic nature of clinical practice, where preventive interventions initiated based on AI predictions could alter the predicted outcomes. For instance, when the model predicts high AKI risk, clinicians might implement preventive measures such as fluid management or discontinuation of nephrotoxic agents, potentially averting the predicted event. This intervention effect needs to be carefully considered during implementation. Additionally, developing user-friendly interfaces that provide intuitive visualizations and clear explanations of AI predictions, and implementing feedback mechanisms for continuous improvement,

are important for successfully integrating AI models such as PRIME Solution into routine clinical practice. The dynamic interplay between AI and human judgment revealed in this study provides valuable insights into the future of healthcare, in which AI and clinicians work synergistically to improve patient outcomes.

Chapter 4. Conclusions

This research explored data-driven approaches in nephrology through two complementary studies: an investigation of environmental factors affecting primary glomerulonephritis and an examination of artificial intelligence applications in acute kidney injury prediction. These studies advance our understanding of kidney diseases through different analytical approaches.

The first study examined links between air pollution exposure and renal function deterioration in primary glomerulonephritis patients. Elevated levels of PM_{10}, $PM_{2.5}$, CO, and NO_2 were associated with increased risk of renal function deterioration, with particulate matter showing the strongest and most consistent effects. These findings highlight the importance of considering environmental factors in chronic kidney disease progression and suggest the value of incorporating air quality monitoring in long-term kidney disease management strategies.

The second study evaluated the effectiveness of artificial intelligence in acute kidney injury prediction through the PRIME Solution model. This AI-based clinical decision support system showed improvements in predictive performance across different levels of medical expertise, particularly in recall and efficiency. The results indicated that while AI assistance enhanced overall prediction accuracy, its impact varied based on clinical expertise

level, with specialists demonstrating the most effective integration of AI recommendations with their clinical judgment.

The combination of these studies adds to the growing body of knowledge in data-driven nephrology. By examining primary glomerulonephritis and acute kidney injury, this research provides insights into kidney disease management through different analytical approaches. These studies demonstrate diverse applications of data science in medical research, from statistical modeling to artificial intelligence-based prediction. Both studies suggest practical implications for patient care, from consideration of environmental risk factors in primary glomerulonephritis to the implementation of early prediction systems for acute kidney injury.

These findings support the potential value of data-driven approaches in nephrology and suggest that future kidney disease management strategies may benefit from incorporating both environmental risk factor monitoring and AI-assisted clinical prediction systems. This integrated approach could contribute to the development of prevention strategies and patient care across the spectrum of kidney diseases.

※ In the preparation of this dissertation, language assistance was provided by AI tools (ChatGPT and Claude). All AI-generated content was thoroughly reviewed and edited, with the author maintaining full responsibility for the entire content of this work.

Bibliography

1. Lameire NH, Levin A, Kellum JA, et al. Harmonizing acute and chronic kidney disease definition and classification: Report of a kidney disease: Improving global outcomes (kdigo) consensus conference. *Kidney Int* 2021;100:516-526.
2. Levey AS. Defining akd: The spectrum of AKI, akd, and ckd. *Nephron* 2022;146:302-305.
3. Collaboration GBDCKD. Global, regional, and national burden of chronic kidney disease, 1990-2017: A systematic analysis for the global burden of disease study 2017. *Lancet* 2020;395:709-733.
4. Kellum JA, Romagnani P, Ashuntantang G, Ronco C, Zarbock A, Anders H-J. Acute kidney injury. *Nat Rev Dis Primers* 2021;7:52.
5. Rajagopalan S, Al-Kindi SG, Brook RD. Air pollution and cardiovascular disease: Jacc state-of-the-art review. *J Am Coll Cardiol*2018. p. 2054-2070.
6. Yang BY, Fan S, Thiering E, et al. Ambient air pollution and diabetes: A systematic review and meta-analysis. *Environ Res* 2020;180:108817.
7. Al-Aly Z, Bowe B. Air pollution and kidney disease. *Clin J Am Soc Nephrol* 2020;15:301-303.
8. Ye JJ, Wang SS, Fang Y, Zhang XJ, Hu CY. Ambient air pollution exposure and risk of chronic kidney disease: A systematic review of the literature and meta-analysis. *Environ Res* 2021;195:110867.
9. Bacci MR, Minczuk CVB, Fonseca FLA. A systematic review of artificial intelligence algorithms for predicting acute kidney injury. *Eur Rev Med Pharmacol Sci* 2023;27:9872-9879.
10. Yi J, Kim SH, Lee H, et al. Air quality and kidney health: Assessing the effects of pm(10), pm(2.5), co, and no(2) on

renal function in primary glomerulonephritis. *Ecotoxicol Environ Saf* 2024;281:116593.
11. Xu C, Zhang Q, Huang G, Huang J, Zhang H. The impact of pm2.5 on kidney. *J Appl Toxicol*: John Wiley and Sons Ltd; 2023. p. 107-121.
12. Chadban SJ, Atkins RC. Glomerulonephritis. *The Lancet* 2005;365:1797-1806.
13. Hu J, Ke R, Teixeira W, et al. Global, regional, and national burden of ckd due to glomerulonephritis from 1990 to 2019 a systematic analysis from the global burden of disease study 2019. *Clin J Am Soc Nephrol* 2023;18:60-71.
14. Glencross DA, Ho TR, Camina N, Hawrylowicz CM, Pfeffer PE. Air pollution and its effects on the immune system. *Free Radic Biol Med* 2020;151:56-68.
15. Zhao CN, Xu Z, Wu GC, et al. Emerging role of air pollution in autoimmune diseases. *Autoimmun Rev* 2019;18:607-614.
16. Bai H, Jiang L, Li T, et al. Acute effects of air pollution on lupus nephritis in patients with systemic lupus erythematosus: A multicenter panel study in china. *Environ Res* 2021;195:110875.
17. Gilcrease GW, Padovan D, Heffler E, et al. Is air pollution affecting the disease activity in patients with systemic lupus erythematosus? State of the art and a systematic literature review. *European Journal of Rheumatology* 2020;7:31-34.
18. Li J, Cui Z, Long J, et al. Primary glomerular nephropathy among hospitalized patients in a national database in china. *Nephrology Dialysis Transplantation* 2018;33:2173-2181.
19. Xu X, Wang G, Chen N, et al. Long-term exposure to air pollution and increased risk of membranous nephropathy in china. *J Am Soc Nephrol* 2016;27:3739-3746.
20. Kwon D, Lee W, Kang C, et al. Estimation of high-spatial resolution of ground-level ozone, nitrogen dioxide, and carbon monoxide in south korea during 2002-2020 using machine-learning based ensemble model. *SSRN Electronic Journal* 2022. DOI:10.2139/ssrn.4280944.

21. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604-612.
22. Bowe B, Xie Y, Li T, Yan Y, Xian H, Al-Aly Z. Particulate matter air pollution and the risk of incident ckd and progression to esrd. *J Am Soc Nephrol* 2018;29:218-230.
23. Luo C, Ouyang Y, Shi S, et al. Particulate matter of air pollution may increase risk of kidney failure in iga nephropathy. *Kidney Int* 2022;102:1382-1391.
24. Miller MR, Raftis JB, Langrish JP, et al. Inhaled nanoparticles accumulate at sites of vascular disease. *ACS Nano* 2017;11:4542-4552.
25. Xiao X, Yao T, Du S, et al. Age differences in the pulmonary and vascular pathophysiologic processes after long-term real-time exposure to particulate matter in rats. *Chemosphere* 2020;261.
26. Xu W, Wang S, Jiang L, et al. The influence of pm(2.5) exposure on kidney diseases. *Hum Exp Toxicol* 2022;41:9603271211069982.
27. Xu MX, Qin YT, Ge CX, et al. Activated irhom2 drives prolonged pm 2.5 exposure-triggered renal injury in nrf2-defective mice. *Nanotoxicology* 2018;12:1045-1067.
28. Kuzma L, Malyszko J, Bachorzewska-Gajewska H, Kralisz P, Dobrzycki S. Exposure to air pollution and renal function. *Sci Rep* 2021;11:11419.
29. Li A, Mei Y, Zhao M, et al. Associations between air pollutant exposure and renal function: A prospective study of older adults without chronic kidney disease. *Environ Pollut* 2021;277.
30. Paoin K, Ueda K, Vathesatogkit P, et al. Long-term air pollution exposure and decreased kidney function: A longitudinal cohort study in bangkok metropolitan region, thailand from 2002 to 2012. *Chemosphere* 2022;287:132117.
31. Oh J, Ye S, Kang DH, Ha E. Association between exposure to fine particulate matter and kidney function: Results from the

korea national health and nutrition examination survey. *Environ Res* 2022;212.
32. Kim H-J, Min J-y, Seo Y-S, Min K-b. Association between exposure to ambient air pollution and renal function in korean adults. *Annals of occupational and environmental medicine* 2018;30:1-7.
33. Lue SH, Wellenius GA, Wilker EH, Mostofsky E, Mittleman MA. Residential proximity to major roadways and renal function. *Journal of Epidemiology and Community Health* 2013;67:629-634.
34. Chen S-Y, Chu D-C, Lee J-H, Yang Y-R, Chan C-C. Traffic-related air pollution associated with chronic kidney disease among elderly residents in taipei city. *Environ Pollut* 2018;234:838-845.
35. Mainka A, Żak M. Synergistic or antagonistic health effects of long- and short-term exposure to ambient no2 and pm2.5: A review. *International Journal of Environmental Research and Public Health*: MDPI; 2022. p. 14079-14079.
36. Mills IC, Atkinson RW, Anderson HR, Maynard RL, Strachan DP. Distinguishing the associations between daily mortality and hospital admissions and nitrogen dioxide from those of particulate matter: A systematic review and meta-analysis. *BMJ Open* 2016;6.
37. Li H, Deng W, Small R, Schwartz J, Liu J, Shi L. Health effects of air pollutant mixtures on overall mortality among the elderly population using bayesian kernel machine regression (bkmr). *Chemosphere* 2022;286.
38. Shumake KL, Sacks JD, Lee JS, Johns DO. Susceptibility of older adults to health effects induced by ambient air pollutants regulated by the european union and the united states. *Aging Clin Exp Res* 2013;25:3-8.
39. Zhang X, Staimer N, Gillen DL, et al. Associations of oxidative stress and inflammatory biomarkers with chemically-

characterized air pollutant exposures in an elderly cohort. *Environ Res* 2016;150:306-319.
40. Wang W, Wu C, Mu Z, et al. Effect of ambient air pollution exposure on renal dysfunction among hospitalized patients in shanghai, china. *Public Health* 2020;181:196-201.
41. Vivarelli M, Massella L, Ruggiero B, Emma F. Minimal change disease. *Clin J Am Soc Nephrol* 2017;12:332-345.
42. Chou YH, Lien YC, Hu FC, et al. Clinical outcomes and predictors for esrd and mortality in primary gn. *Clin J Am Soc Nephrol* 2012;7:1401-1408.
43. Ji JS, Liu L, Zhang JJ, et al. No(2) and pm(2.5) air pollution co-exposure and temperature effect modification on premature mortality in advanced age: A longitudinal cohort study in china. *Environ Health* 2022;21:97.
44. Jeong S-M, Park J-H, Kim H-J, Kwon H, Hwang SE. Effects of abdominal obesity on the association between air pollution and kidney function. *International Journal of Obesity* 2020;44:1568-1576.
45. Li L, Zhang W, Liu S, et al. Associations of multiple air pollutants with kidney function in normal-weight and obese adults and effect modification by free fatty acids. *Chemosphere* 2023;341.
46. Park S, Baek SH, Ahn S, et al. Impact of electronic acute kidney injury (AKI) alerts with automated nephrologist consultation on detection and severity of AKI: A quality improvement study. *Am J Kidney Dis* 2018;71:9-19.
47. Ramgopal S, Sanchez-Pinto LN, Horvat CM, Carroll MS, Luo Y, Florin TA. Artificial intelligence-based clinical decision support in pediatrics. *Pediatr Res* 2023;93:334-341.
48. Nagendran M, Festor P, Komorowski M, Gordon AC, Faisal AA. Quantifying the impact of AI recommendations with explanations on prescription decision making. *NPJ Digit Med* 2023;6:206.

49. Jeon H, Jang HR. Electronic alerts based on clinical decision support system for post-contrast acute kidney injury. *Kidney Res Clin Pract* 2023;42:541-545.
50. Zhang H, Wang AY, Wu S, et al. Artificial intelligence for the prediction of acute kidney injury during the perioperative period: Systematic review and meta-analysis of diagnostic test accuracy. *BMC Nephrol* 2022;23:405.
51. Bajgain B, Lorenzetti D, Lee J, Sauro K. Determinants of implementing artificial intelligence-based clinical decision support tools in healthcare: A scoping review protocol. *BMJ Open* 2023;13:e068373.
52. Kim K, Yang H, Yi J, et al. Real-time clinical decision support based on recurrent neural networks for in-hospital acute kidney injury: External validation and model interpretation. *J Med Internet Res* 2021;23:e24120.
53. Bach S, Binder A, Montavon G, Klauschen F, Muller KR, Samek W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS One* 2015;10:e0130140.
54. Patterson ES, Doebbeling BN, Fung CH, Militello L, Anders S, Asch SM. Identifying barriers to the effective use of clinical reminders: Bootstrapping multiple methods. *J Biomed Inform* 2005;38:189-199.
55. Khwaja A. Kdigo clinical practice guidelines for acute kidney injury. *Nephron Clin Pract* 2012;120:c179-184.
56. Alzubaidi L, Zhang J, Humaidi AJ, et al. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 2021;8:1-74.
57. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 2018.
58. Yang M, Liu S, Hao T, et al. Development and validation of a deep interpretable network for continuous acute kidney injury

prediction in critically ill patients. *Artif Intell Med* 2024;149:102785.
59. Mistry NS, Koyner JL. Artificial intelligence in acute kidney injury: From static to dynamic models. *Adv Chronic Kidney Dis* 2021;28:74-82.
60. Bell S, James MT, Farmer CK, Tan Z, De Souza N, Witham MD. Development and external validation of an acute kidney injury risk score for use in the general population. *Clin Kidney J* 2020;13:402-412.
61. Schwager E, Ghosh E, Eshelman L, Pasupathy KS, Barreto EF, Kashani K. Accurate and interpretable prediction of ICU-acquired AKI. *J Crit Care* 2023;75:154278.
62. Castelvecchi D. Can we open the black box of AI? *Nature* 2016;538:20-23.
63. Flechet M, Falini S, Bonetti C, et al. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: A prospective evaluation of the AKIpredictor. *Crit Care* 2019;23:1-10.
64. Inkpen K, Chappidi S, Mallari K, et al. Advancing human-AI complementarity: The impact of user expertise and algorithmic tuning on joint decision making. *ACM Trans Comput Hum Interact* 2023;30:1-29.
65. Wang X, Yin M. Are explanations helpful? A comparative study of the effects of explanations in ai-assisted decision-making. *Proceedings of the 26th International Conference on Intelligent User Interfaces* 2021:318-328.
66. Shamszare H, Choudhury A. Clinicians' perceptions of artificial intelligence: Focus on workload, risk, trust, clinical decision making, and clinical integration. *Healthcare* 2023;11:2308.
67. Glick A, Clayton M, Angelov N, Chang J. Impact of explainable artificial intelligence assistance on clinical decision-making of novice dental clinicians. *JAMIA open* 2022;5:ooac031.

68. Khosravi M, Zare Z, Mojtabaeian SM, Izadi R. Artificial intelligence and decision-making in healthcare: A thematic analysis of a systematic review of reviews. *Health Serv Res Manag Epidemiol* 2024;11:23333928241234863.
69. Khalifa M, Albadawy M, Iqbal U. Advancing clinical decision support: The role of artificial intelligence across six domains. *Comput Methods Programs Biomed Update* 2024;5:100142.
70. Dubois C, Le Ny J. Adaptive task allocation in human-machine teams with trust and workload cognitive models. *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* 2020:3241-3246.
71. Lee Y, Kim T, Kim DE, et al. Differences in the incidence, characteristics, and outcomes of patients with acute kidney injury in the medical and surgical intensive care units. *Kidney Res Clin Pract* 2024;43:518-527.
72. Cheungpasitporn W, Thongprayoon C, Kashani KB. Artificial intelligence and machine learning's role in sepsis-associated acute kidney injury. *Kidney Res Clin Pract* 2024;43:417-432.
73. Choi H, Choi B, Han S, et al. Applicable machine learning model for predicting contrast-induced nephropathy based on pre-catheterization variables. *Internal Medicine* 2024;63:773-780.

Abstract in Korean

급성 및 만성 신장질환은 현대 의학과 전 세계 보건의료 시스템이 직면한 주요 과제이다. 신장 기능 저하의 진행에는 다양한 병리학적 요인이 기여할 수 있어, 이러한 질환들의 효과적인 관리를 위해서는 다양한 위험 요인에 대한 이해와 함께 견고한 예측 전략의 개발이 필요하다. 신장학 분야에서 시계열 데이터의 중요성이 점차 부각되고 있으며, 이는 질병 진행 추적과 예방 전략 수립에 핵심적인 역할을 한다. 고급 분석 기법과 인공지능을 활용한 시계열 데이터의 심층 분석은 신장질환의 예측과 모니터링에 혁신적인 접근 방식을 제공하여, 환자 진료의 질과 의료 시스템의 효율성을 향상시킬 수 있는 잠재력을 보여준다. 본 연구는 두 가지 연구를 결합했다: 하나는 일차성 사구체신염에서 신장 기능에 영향을 미치는 환경적 요인을 조사한 연구이고, 다른 하나는 급성 신장 손상 예측에 기계학습을 적용한 연구로, 이를 통해 다양한 유형의 신장 질환에서 데이터 기반 접근법에 대한 포괄적인 관점을 제시한다.

(1) 대기질과 신장 건강: 일차성 사구체신염에서 PM_{10}, $PM_{2.5}$, CO, NO_2가 신장 기능에 미치는 영향 평가

연구 배경: 대기오염 노출과 심혈관 질환, 당뇨병과 같은 만성질환 간의 관계는 광범위하게 연구되어 왔으나, 면역 매개성 신장 질환인 일차성 사구체신염(GN)에 대한 구체적인 영향은 아직 충분히 이해되지 않았다. GN의 발생률이 증가하고 대기질과의 연관성에 대한 연구가 부족한 점을 고려하여, 일차성 GN 환자의 신장 기능에 대한 대기오염 물질의 장기적 영향을 조사했다.

연구 방법: 서울대학교병원과 분당서울대학교병원에서 진단받은 일

차성 GN 환자 1,394명을 대상으로 후향적 코호트 분석을 수행했다. 시변 Cox 회귀분석과 선형혼합모형(LMM)을 사용하여 연평균 대기오염 수준이 신장 기능 저하(RFD)와 추정 사구체 여과율(eGFR) 변화에 미치는 영향을 분석했다. 여기서 RFD는 지속적으로 eGFR이 60 mL/min per 1.73 m^2 미만인 상태로 정의했다.

연구 결과: 평균 5.1년의 관찰 기간 동안 350명의 참가자가 RFD를 보였다. 인구통계학적 변수와 건강 변수를 고려한 후에도, 대기오염 물질의 사분위수 범위(IQR) 수준 증가—PM10(직경 10마이크로미터 이하 입자, HR 1.389, 95% CI 1.2−1.606), PM2.5(직경 2.5마이크로미터 이하 입자, HR 1.353, 95% CI 1.162−1.575), CO(일산화탄소, HR 1.264, 95% CI 1.102−1.451), NO2(이산화질소, HR 1.179, 95% CI 1.021−1.361)—는 RFD 위험 증가와 유의한 연관성을 보였다. 또한 PM10, PM2.5, CO 노출은 eGFR 감소와 연관성을 보였다.

결론: 본 연구는 일차성 GN에서 대기오염 노출과 신장 기능 손상 간의 실질적인 연관성을 입증했으며, 면역 매개성 신장질환의 병리에서 환경적 결정 요인의 중요성을 강조한다.

(2) 임상 의사 결정 지원 시스템으로서의 급성 신장 손상 예측 모델 검증

연구 배경: 급성 신장 손상(AKI)은 즉각적인 중재가 필요한 중요한 임상 상태이다. 본 연구진은 AKI를 예측하기 위한 PRIME Solution이라는 인공지능(AI) 모델을 개발하고, 임상의의 예측 능력 향상에 대한 기여도를 평가했다.

연구 방법: PRIME Solution은 3차 의료기관의 183,221건의 입원

기록(2013–2017)을 사용하여 잔차 블록이 있는 합성곱 신경망으로 개발되었으며, 다른 3차 의료기관의 4,501건의 입원 기록(2020–2021)으로 외부 검증을 실시했다. 응용 평가를 위해 후자 병원의 환자 100명(AKI 사례 15건 포함)의 후향적 수집 데이터를 사용하여 전향적 평가를 수행했다. AKI 예측 성능은 전문의, 일반의, 의과대학생 간에 AI 지원 유무에 따라 비교했다.

연구 결과: AI 지원 없이는 전문의가 가장 높은 정확도(0.797)를 보였고, 그 다음으로 의과대학생(0.619)과 PRIME Solution(0.568) 순이었다. AI 지원은 전반적인 재현율(61.0%에서 74.0%)과 F1 점수(38.7%에서 42.0%)를 향상시켰으며, 평균 검토 시간을 단축시켰다(73.8초에서 65.4초; $p<0.001$). 그러나 전문성 수준에 따라 그 영향은 달랐다. 전문의가 가장 큰 향상을 보였고(재현율: 32.1%에서 64.3%; F1: 36.4%에서 48.6%), 의과대학생의 성능은 향상되었으나 AI 모델과 더 유사한 수준으로 수렴했다. 또한 AI 지원의 효과는 예측 결과에 따라 달랐는데, AKI로 예측된 사례에서는 재현율이, 비AKI로 예측된 사례에서는 정밀도, F1 점수, 검토 시간 단축(73.4초에서 62.1초; $p<0.001$)이 더 큰 향상을 보였다.

결론: AI 지원으로 AKI 예측이 향상되었으나, 그 개선 정도는 사용자의 전문성에 따라 차이를 보였다.

결론적으로, 이 연구들은 신장학 분야에서 데이터 기반 접근법의 잠재력을 강조한다. 첫 번째 연구는 대기오염과 만성 신장 기능 저하 간의 명확한 연관성을 확립하여 장기적인 신장질환 관리에서 환경적 요인 고려의 중요성을 강조했다. 두 번째 연구는 급성 신장 손상 예측에서 AI의

가치를 입증하여 신속한 중재를 위한 임상 실무에서 고급 분석의 잠재력을 보여주었다. 이러한 발견들은 급성 손상에서 만성 질환에 이르기까지 신장질환 전반에 걸쳐 환자의 예후와 삶의 질을 개선할 수 있는 더욱 표적화된 중재와 개별화된 치료 전략의 길을 열어준다.

주제어: 만성 신장질환, 급성 신장 손상, 대기 오염, 기계 학습, 시계열 분석, 사구체신염

학 번: 2021-33712