



Article An Approach for Personalized Dynamic Assessment of Chronic Kidney Disease Progression Using Joint Model

Chen-Mao Liao ^{1,†}, Yi-Wei Kao ^{1,†}, Yi-Ping Chang ^{2,3} and Chih-Ming Lin ^{3,*}

- ¹ Department of Applied Statistics and Information Science, Ming Chuan University, Taoyuan 333, Taiwan; cmliao@mail.mcu.edu.tw (C.-M.L.); ywkao@mail.mcu.edu.tw (Y.-W.K.)
- ² Department of Nephrology, Taoyuan Branch of Taipei Veterans General Hospital, Taoyuan 333, Taiwan; ypchangtyvh@gmail.com
- ³ Department of Healthcare Information and Management, Ming Chuan University, Taoyuan 333, Taiwan
- Correspondence: cmlin@mail.mcu.edu.tw; Tel.: +886-3-350-7001 (ext. 3530); Fax: +886-3-359-3880
- [†] These authors contributed equally to this work.

Abstract: Chronic kidney disease (CKD) poses significant challenges to public health and healthcare systems, demanding a comprehensive understanding of its progressive nature. Prior methods have often fallen short in capturing the dynamic and individual variability of renal function. This study aims to address this gap by introducing a novel approach for the individualized assessment of CKD progression. A cohort of 1042 patients, comprising 700 with stage 3a and 342 with stage 3b to stage 5 CKD, treated at a veteran general hospital in Taiwan from 2006 to 2019, was included in the study. A comprehensive dataset spanning 12 years, consisting of clinical measurements, was collected and analyzed using joint models to predict the progression to hemodialysis treatment. The study reveals that the estimated glomerular filtration rate (eGFR) can be considered an endogenous factor influenced by innate biochemical markers. Serum creatinine, blood pressure, and urinary protein excretion emerged as valuable factors for predicting CKD progression. The joint model, combining longitudinal and survival analyses, demonstrated predictive versatility across various CKD severities. This innovative approach enhances conventional models by concurrently incorporating both longitudinal and survival analyses and provides a nuanced understanding of the variables influencing renal function in CKD patients. This personalized model enables a more precise assessment of renal failure risk, tailored to each patient's unique clinical profile. The findings contribute to improving the management of CKD patients and provide a foundation for personalized healthcare interventions in the context of renal diseases.

Keywords: chronic kidney disease; hemodialysis; glomerular filtration rate

1. Introduction

Chronic kidney disease (CKD) is a progressive, worsening disease, and the increasing global prevalence of CKD presents substantial challenges to public health and healthcare systems internationally. There are approximately 11% of the world's 7.5 billion population, which means that more than 800 million people are affected by CKD [1,2]. On the other hand, CKD is also widely recognized as one of the leading causes of death in the world's population at present, with global all-age mortality due to CKD increasing by more than 40% between 1990 and 2017 [3]. Ranked 12th in 2017 compared to 36th in 1990, with projections indicating that it will become the fifth leading cause of life lost globally by 2040 [2,4]. In the United States, approximately 8 million adults are suffering from CKD with at least stage 3, while more than 400,000 individuals have end-stage renal disease (ESRD), leading to significant medical expenditures and increased mortality rates, particularly cardiovascular complications [3,5–7]. It is worth noting that Taiwan exhibits a higher incidence and prevalence of CKD compared to other regions and the highest prevalence of ESRD in the world, rendering it an area of critical concern [5,8,9].



Citation: Liao, C.-M.; Kao, Y.-W.; Chang, Y.-P.; Lin, C.-M. An Approach for Personalized Dynamic Assessment of Chronic Kidney Disease Progression Using Joint Model. *Biomedicines* 2024, *12*, 622. https://doi.org/10.3390/ biomedicines12030622

Academic Editors: Shuo-Ming Ou and Shang-Feng Yang

Received: 18 February 2024 Revised: 4 March 2024 Accepted: 8 March 2024 Published: 11 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). Given this global challenge, understanding the mechanisms of CKD progression is critical. The progression of CKD is closely associated with the decline in renal function, especially the reduction in the estimated glomerular filtration rate (eGFR) [10]. This decline not only indicates progression to more advanced stages of CKD but also potential toward end-stage renal disease (ESRD) [10]. Therefore, early identification and management of eGFR decline are imperative in decelerating the progression of CKD to ESRD. Recent technological advancements, especially the integration of machine learning in CKD assessment, have facilitated novel methodologies for the early prediction and management of the disease [11–13]. These advancements highlight the criticality of early detection and precise prognostication in understanding the progression of CKD. However, inadequate capturing of the dynamic and individual variability of renal function over time remained in the prediction methods [14,15].

To overcome these limitations, a more sophisticated approach is required. Assembled with the advantage of both the Cox model and the mixture model, the joint model offers an advanced methodology for predicting individual physiological function trajectories and related health outcomes. It facilitates the analysis of both static attributes (e.g., gender) and dynamic characteristics (e.g., age), adeptly addressing challenges such as intermittent data collection, measurement inaccuracies, and the truncation of measurement processes associated with survival risks. However, the application of this method in the field of CKD, particularly in individualized assessment of renal function failure using eGFR, is still relatively limited [16].

To fill this gap, we have developed a methodological framework centered around these innovative models. The purpose of this study is to develop and apply a novel joint model using eGFR for individualized assessment of CKD progression. This model will account for the nonlinear dynamics of renal function changes and simultaneously handle multiple time-to-event outcomes related to renal function decline. This will provide more accurate and individualized prognostic assessments for patients with CKD, potentially improving clinical decision-making and patient management.

2. Methods

2.1. Study Design and Patient Population

In this retrospective cohort study, we aimed to assess the progression of chronic kidney disease (CKD) in patients at stages 3 to 5 who are characterized by an eGFR of less than 60 mL/min/1.73 m². Data spanning from January 2006 to July 2019 were obtained from the Taoyuan Branch of Taipei Veterans General Hospital. The analysis predominantly utilized data from two National Health Insurance Administration (NHIA) initiatives: the 'NHI Pre-ESRD Patient Care and Education Program' and the 'NHI Reimbursement Plans that Improve Health Care Quality of Early-Stage Chronic Kidney Disease'. These initiatives were instrumental in providing the essential data for this study.

The eligibility criteria for patients with chronic kidney disease (CKD) encompassed individuals who received their initial treatment for CKD at the hospital and subsequently participated in follow-up sessions occurring at least twice a year. Diagnosis confirmation was strictly based on the ICD-9-CM and ICD-10 coding systems chosen for their comprehensive coverage of CKD-related conditions. The study's methodology encompassed a comprehensive approach involving regular health assessments, including laboratory tests and physical measurements, supplemented by educational sessions about CKD. Participants in the early CKD intervention group were assessed biannually, whereas those in the advanced pre-ESRD care group received assessments at least once every three months, continuing until the initiation of dialysis or renal transplantation as of 31 July 2019.

Patients who transferred to other medical facilities, passed away, or withdrew from the study for various reasons unrelated to renal disease during this period were classified under non-ESRD endpoints. The hazard of evolving into ESRD was calculated in 90-day intervals from the patient's initial nephrology consultation to the end of the follow-up. Following the exclusion of 24 individuals due to incomplete records, where missing data exceeded 30%, the remaining participants were categorized based on their enrollment in the aforementioned NHIA programs. The final cohort included 1042 patients, categorized into 700 with stage 3a, 111 with stage 3b, 143 with stage 4, and 88 with stage 5 CKD. The count of patients per CKD stage commencing dialysis therapy during the study was 48 for stage 3a, 2 for stage 3b, 17 for stage 4, and 40 for stage 5.

2.2. Variables

This study primarily focused on ESRD, characterized by the initiation of either hemodialysis or peritoneal dialysis treatments. It is important to note that none of the subjects in this study underwent kidney transplantation. Data for this study, spanning from January 2006 to July 2019, were sourced from the study hospital. This study utilized information from two NHIA programs, including the 'NHI Pre-ESRD patient care and education program' and the 'NHI reimbursement plans that enhance the quality of care for early-stage chronic kidney disease.'

The clinical characteristics of the subjects were organized into four main categories: (1) Demographic factors: gender, age, height, and weight. (2) Laboratory measurements indicative of CKD severity: eGFR, hemoglobin (Hgb), hematocrit (Hct), serum albumin (Alb), creatinine (Cr), blood urea nitrogen (BUN), sodium (Na), potassium (K), calcium (Ca), phosphorus (P), triglyceride (Tri), and urine protein–creatinine ratio (PCR). (3) Comorbid conditions: hypertension, diabetes, and cardiovascular diseases (CVDs). (4) Additional risk-related biophysical and biochemical markers: blood pressure, uric acid, lipid profiles, fasting glucose, and HbA1c levels.

The eGFR was calculated using the simplified modification of diet in renal disease (MDRD) equation as follows [17]: eGFR = $186 \times \text{age}^{-}-0.203 \times \text{Cr}^{-}-1.154 \times (0.742 \text{ if female})$. Particularly validated for accuracy in patients with CKD stages 3 to 5, this formula demonstrates superiority when compared to the Cockcroft–Gault (CG) and the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equations [18,19].

During the study, follow-up procedures and biochemistry testing varied between the two NHIA programs, with blood tests being conducted during clinic visits according to patient-specific designated protocols. Specifically, follow-up intervals were set every 6 months for stage 3a CKD patients and at least every 3 months for patients in stages 3b to 5, continuing until death, the initiation of dialysis, or loss to follow-up. Baseline characteristics and laboratory variables were initially derived from the first clinic visit, with only the most recent measurements included in the dynamic analysis within each subsequent 90-day period to ensure data consistency.

For stage 3a CKD patients, a narrower set of clinical characteristics was analyzed due to variations in the available laboratory datasets between the two NHIA programs. Parameters like Hgb, Hct, ALB, BUN, Na, K, Ca, P, uric acid, cholesterol, triglycerides, and fasting glucose were omitted. Variables with more than 30% missing values were excluded. For the remaining variables with incomplete data, multiple imputation was performed using the multivariate imputation by chained equations (MICE) module in the R software package [20].

2.3. Statistical Analysis

In this study, the joint model was utilized to analyze patient data across various stages of CKD. We focused predominantly on two models: one dedicated to CKD stage 3a and another encompassing stages 3b to 5.

Each joint model consisted of a linear mixed-effects model (LME) component, incorporating clinically relevant variables like observation time (obstime), systolic blood pressure (SBP), diastolic blood pressure (DBP), creatinine, low-density lipoprotein cholesterol (LDL), glycated hemoglobin (HbA1c), the logarithm of protein-to-creatinine ratio (PCRln), age, sex, hypertension, and diabetes. These variables were selected based on their clinical significance in the progression of CKD. To clarify the relationships between various factors and CKD progression, our analysis incorporated two principal statistical assessments. Independent sample t-tests were applied to continuous variables, including GFR, height, weight, SBP, DBP, Cr, LDL, HbA1c, PCRIn, and age, facilitating the determination of *p*-values to assess the statistical significance of differences across distinct groups. This methodology enabled a comprehensive evaluation of the contribution of each continuous variable to disease dynamics.

On the other hand, categorical variables such as sex, hypertension, diabetes, and cardiovascular disease (CVD) were analyzed using chi-square tests of independence to calculate *p*-values. This method helped us identify whether there are significant associations between these categorical variables and the progression or stages of CKD.

Simultaneously, the Cox proportional hazards model component was employed to quantify the dialysis risk. This model encompassed variables including creatinine, proteinto-creatinine ratio, age, sex, hypertension, and diabetes. The choice of variables in our Cox model was based on both statistical and clinical grounds, aiming for the lowest Akaike Information Criterion (AIC). This approach underscores our dedication to achieving a balance between model parsimony and predictive accuracy in determining dialysis risk in CKD patients. The models were cohesively fitted using the jointModel function in R, enabling the concurrent analysis of time-to-event data and continuous outcomes.

In this study, the progression of renal function and the risk of renal failure in patients with CKD were analyzed using a joint modeling approach. The model comprises two main processes: the longitudinal process and the event process.

3. Longitudinal Process

The longitudinal process is typically described using a linear mixed-effects model (LME). For our case, the model is expressed as follows:

$$GFR_{ij} = \beta_0 + \sum_{k=1}^{p} \beta_k X_{kij} + \sum_{l=1}^{q} b_{il} Z_{lij} + \epsilon_{ij}$$

where:

- *GFR*_{*ij*} is the glomerular filtration rate for the *i*th patient at *j*th time point.
- $\beta_0, \beta_1, \ldots, \beta_p$ are the fixed effect parameters.
- *X_{kij}* are observed covariates (like SBP, creatinine, age).
- b_{i0}, \ldots, b_{iq} are the random effect parameters for each patient.
- *Z*_{*lij*} are variables associated with the random effects.
- ϵ_{ii} is the random error term.

4. Event Process

The event process is modeled using a Cox proportional hazards model, expressed as follows:

$$h(t_i) = h_0(t) \exp\left(\sum_{m=1}^r \gamma_m W_{mi} + \alpha f(GFR_i)\right)$$

where:

- $h(t_i)$ is the hazard function for the *i*th patient at time *t*.
- $h_0(t)$ is the baseline hazard function.
- $\gamma_1, \ldots, \gamma_r$ are fixed effect coefficients for the covariates.
- *W_{mi}* are the covariates in the event process (like proteinuria, diabetes).
- *α* is the association parameter, representing the link between the longitudinal and event processes.
- *f*(*GFR_i*) is a function of the GFR over time in the longitudinal process.

During the data analysis, the selection of models was rigorously evaluated and guided by the Akaike Information Criterion (AIC), a measure used to determine the model's fit while penalizing complexity. This criterion was instrumental in identifying the two most effective models for elucidating the relationship between the decline in GFR and the escalating risk of dialysis in CKD patients. The applicability of these selected models encompassed assessments of model stability and predictive accuracy through specific statistical tests, as well as a detailed inspection of model parameters to ensure their scientific validity and rationality.

5. Results

In Table 1 of our study, we noted significant differences in GFR, serum creatinine levels, and PCRIn between patients requiring dialysis and those not requiring it in stage 3a chronic kidney disease. All these differences had *p*-values less than 0.001. This observation indicates the impact of different treatment approaches on the physiological state of patients. However, no statistically significant differences were found in height, weight, systolic blood pressure, and diastolic blood pressure between the two groups. This suggests a consistency in these fundamental physical parameters across different treatment modalities.

Variable Non-Dialysis Dialysis *p*-Value Mean (SD)/n (%) Mean (SD)/n (%) $GFR (mL/min/1.73 m^2)$ 53.92 (10.21) 46.61 (10.83) < 0.001Height (cm) 160.52 (11.45) 161.15 (7.89) 0.714 Weight (kg) 65.88 (12.43) 64.59 (12.54) 0.505 SBP (mmHg) 132.86 (17.94) 137.48 (19.00) 0.098 DBP (mmHg) 72.82 (10.90) 71.30 (11.22) 0.371 1.30 (0.24) Cr (mg/dL)1.52 (0.42) < 0.001 LDL (mg/dL) 102.39 (28.16) 96.50 (29.71) 0.178 6.88 (1.72) 0.098 HbA1c (%) 6.52 (1.35) 5.23 (1.24) 6.11 (1.54) < 0.001 PCRln (mg/g)80.74 (11.30) 79.57 (13.09) 0.507 Age (year) Sex 0.846 Female 144 (30.8%) 13 (28.3%) Male 323 (69.2%) 33 (71.7%) Hypertension 0.985 Yes 189 (40.5%) 18 (39.1%) 278 (59.5%) No 28 (60.9%) 0.024 Diabetes 229 (49.0%) 14 (30.4%) Yes No 238 (51.0%) 32 (69.6%) CVD 0.271 357 (76.4%) 39 (84.8%) Yes 110 (23.6%) 7 (15.2%) No 3.77 (1.70) 3.30 (1.74) 0.080 Frequency

Table 1. Patient characteristics of group 3a.

In our study focusing on stage 3a chronic kidney disease (CKD) patients, the application of a joint model revealed detailed insights into the progression toward renal failure, as shown in Table 2.

For these patients, observational time (obstime) was a critical factor, indicating a significant time-dependent decline in renal function. Systolic blood pressure (SBP) emerged as a considerable predictor (estimate: 0.046, *p*-value: 0.0005), suggesting that higher SBP levels correlate with an accelerated deterioration in renal function. Serum creatinine, a pivotal biomarker, displayed an inverse relationship with renal health (estimate: -2.096, *p*-value: <0.0001), reinforcing its utility in monitoring CKD progression. Age also played a significant role (estimate: -0.088, *p*-value: 0.0167), with older patients showing a more rapid decline.

Model Component	Parameter	Estimate	Std. Error	<i>p</i> -Value
Longitudinal Process	(Intercept)	57.122	3.484	< 0.0001
5	Obstime	-0.002	0.001	0.0002
	SBP	0.046	0.013	0.0005
	Cr	-2.096	0.210	< 0.0001
	Age	-0.088	0.037	0.0167
Event Process	PCRln	0.604	0.146	< 0.0001
	Male	0.933	0.382	0.0146
	Diabetes	0.060	0.350	0.8637
	Hypertension	0.531	0.327	0.1042
	Assoct	-0.117	0.027	< 0.0001
	log(xi.1)	-9.104	1.824	
	log(xi.2)	-8.044	1.796	
	log(xi.3)	-7.732	1.785	
	log(xi.4)	-8.961	1.827	
	log(xi.5)	-7.576	1.701	
	log(xi.6)	-8.121	1.729	
	log(xi.7)	-6.695	1.664	

Table 2. Joint model results for the group 3a.

PCRln levels were strongly associated with an increased risk of renal failure (estimate: 0.604, *p*-value: <0.0001), confirming its vital role in managing CKD. The gender factor, particularly being male, was identified as a risk factor (estimate: 0.933, *p*-value: 0.0146), suggesting the need for gender-specific approaches in CKD treatment. Contrary to initial expectations, diabetes did not emerge as a strong predictor, indicating that other factors might be more influential in the progression of CKD in our patient cohort. Hypertension showed a significant, independent impact on renal failure risk (estimate: 0.531, *p*-value: 0.1042), highlighting the necessity of effective blood pressure management. The association constant (Assoct) successfully linked the longitudinal changes in renal function with the risk of renal failure (estimate: -0.117, *p*-value: <0.0001), demonstrating the complex interplay of these aspects in CKD progression.

Figures 1 and 2 exemplify the predictive power of the joint modeling approach for patients at a similar stage of CKD but with divergent prognoses regarding the onset of dialysis.

Figure 1 portrays the predictive trajectory for Subject 256, a patient in stage 3a CKD, who is characterized by a later transition to dialysis. The depicted median survival curve, supported by the 95% confidence intervals, illustrates a gradual decline in the non-dialysis probability over a span of 3650 days. The finer resolution at selected follow-up intervals—days 1442, 1799, 2149, and 2513—reveals a consistently moderate descent in the survival probability, indicating a delayed progression toward end-stage renal disease.

Figure 2 contrasts this with Subject 647, also in stage 3a CKD but displaying an earlier requirement for dialysis intervention. The patient's median survival curve descends more steeply, suggesting a more rapid decline in renal function. This is corroborated by the individual plots at follow-up times of 721, 1085, 1436, and 1793 days, where each subsequent survival probability estimate shows a pronounced decrease.

The comparative analysis of Figures 1 and 2 underscores the heterogeneity of CKD progression within the same clinical stage. Despite sharing a diagnosis of stage 3a CKD, Subject 647's data indicate a more aggressive course of disease progression, necessitating earlier dialytic support than Subject 256. This divergence accentuates the joint model's ability to provide personalized and dynamic assessments, which are crucial for the anticipatory clinical management of CKD patients.

In Table 3, patients with stage 3b to 5 CKD on dialysis exhibited significantly lower GFR and higher creatinine levels compared to those not on dialysis, reflecting more severe kidney impairment. Differences were also present in metabolic parameters such as urea



nitrogen, phosphorus, and albumin. Blood pressure, lipid profiles, and glucose control, as well as the incidence of common comorbidities, were similar across both groups.

Figure 1. Stage 3a CKD example: late dialysis with the joint model. **(A)** Displays days since observation starts on the horizontal axis (0 to over 3000 days) and the probability of not undergoing dialysis on the vertical axis (0 to 1.0). The solid blue line represents the estimated probability of non-dialysis for case number 256, decreasing over time. Dashed black lines show the 95% confidence interval, widening over time to indicate increasing uncertainty in long-term forecasts. **(B–E)** Mark follow-up time at the top, with the solid green line for predicted GFR changes and dashed black lines for the 95% confidence interval of non-dialysis probability, indicating greater uncertainty over time. Asterisks (*) denote actual observed GFR values.



Figure 2. Stage 3a CKD example: early dialysis with the joint model. (**A**) Displays days since observation starts on the horizontal axis (0 to over 3000 days) and the probability of not undergoing dialysis on the vertical axis (0 to 1.0). The solid blue line represents the estimated probability of non-dialysis for case number 647, decreasing over time. Dashed black lines show the 95% confidence interval, widening over time to indicate increasing uncertainty in long-term forecasts. (**B**–**E**) Mark follow-up time at the top, with the solid green line for predicted GFR changes and dashed black lines for the 95% confidence interval of non-dialysis probability, indicating greater uncertainty over time. Asterisks (*) denote actual observed GFR values.

Variable	Non-Dialysis	Dialysis	<i>p</i> -Value
	Mean (SD)/ <i>n</i> (%)	Mean (SD)/ <i>n</i> (%)	
GFR (mL/min/1.73 m ²)	26.94 (10.74)	11.81 (7.27)	< 0.001
SBP (mmHg)	137.72 (19.35)	141.92 (21.98)	0.142
DBP (mmHg)	74.37 (32.27)	75.44 (14.21)	0.802
Hgb (mg/dL)	11.42 (2.05)	9.62 (1.60)	< 0.001
Hct (%)	34.63 (5.96)	29.89 (5.33)	< 0.001
BUN (mg/dL)	41.44 (22.46)	78.01 (41.92)	< 0.001
Cr (mg/dL)	2.68 (1.57)	6.22 (3.40)	< 0.001
Na (mg/dL)	140.18 (3.48)	139.77 (5.05)	0.452
Ca (mg/dL)	8.92 (0.60)	8.31 (0.80)	< 0.001
P(mg/dL)	3.96 (0.83)	5.06 (1.51)	< 0.001
Alb (g/dL)	3.62 (0.46)	3.16 (0.63)	< 0.001
Tri (mg/dL)	134.71 (79.79)	159.64 (100.05)	0.039
LDL (mg/dL)	102.32 (33.06)	104.68 (37.91)	0.628
HbA1c (%)	6.50 (1.40)	6.66 (1.37)	0.423
PCRln (mg/g)	6.33 (1.76)	7.69 (1.63)	< 0.001
Age (year)	80.25 (13.20)	74.97 (13.55)	0.006
Sex			0.615
Female	88 (32.8%)	22 (37.3%)	
Male	180 (67.2%)	37 (62.7%)	
Hypertension			0.141
Yes	137 (51.1%)	37 (62.7%)	
No	131 (48.9%)	22 (37.3%)	
Diabetes			0.251
Yes	196 (73.1%)	48 (81.4%)	
No	72 (26.9%)	11 (18.6%)	
CVD			0.353
Yes	21 (7.8%)	2 (3.4%)	
No	247 (92.2%)	57 (96.6%)	
Frequency	3.78 (3.20)	4.02 (3.66)	0.611

Table 3. Patient characteristics of groups 3b to 5.

In our study focusing on patients with stage 3b to 5 CKD, as shown in Table 4, we established a joint model that integrates a longitudinal process with an event process to predict renal failure. The longitudinal component revealed that systolic blood pressure (SBP) with an estimate of -0.021 (p < 0.0001) and diastolic blood pressure (DBP) with 0.007 (p = 0.0083) play significant roles in kidney function, demonstrating the critical impact of blood pressure management in CKD progression. Notably, higher serum creatinine levels (estimate: -2.563, p < 0.0001) and phosphate levels (estimate: -0.014, p < 0.0001) were associated with poorer renal outcomes, underscoring their importance as markers of kidney health. Age (estimate: -0.056, p < 0.0001) and proteinuria, as indicated by PCRIn (estimate: -0.775, p < 0.0001), also emerged as crucial factors negatively impacting renal function.

In the event process of the model, Male showed a significant positive estimate of 0.339 (p < 0.0001), suggesting a higher risk of renal failure in males compared to females. Diabetes (estimate: 0.304, p < 0.0001) and hypertension, with an estimate of 0.244 (p < 0.0001), were also identified as significant predictors of renal failure. The association constant (Assoct) presented a negative estimate of -0.153 (p < 0.0001), indicating a complex interaction between the progression of CKD and the risk factors.

Figures 3 and 4 extend the examination of the joint model's predictions for renal failure in patients with more advanced CKD, covering stages 3b to 5.

Figure 3 presents the case of Subject 207, illustrating a late dialysis onset in patients with stages 3b to 5 CKD. The main plot indicates a relatively stable non-dialysis probability over the 3650-day period, with the median survival curve only gradually descending, as delineated by the 95% confidence intervals. This gradual trajectory suggests a slower progression toward renal failure. The individual subplots at follow-up times (727, 846, 930,

and 1021 days) offer a closer look at the predicted probabilities over time, supporting the main plot's indication of delayed dialysis requirement.

Model Component	Parameter	Estimate	Standard Error	<i>p</i> -Value
Longitudinal Process	(Intercept)	37.385	< 0.001	< 0.0001
0	obstime	< 0.001	0.001	0.4422
	SBP	-0.021	0.001	< 0.0001
	DBP	0.007	0.003	0.0083
	Cr	-2.563	< 0.001	< 0.0001
	Р	-0.014	< 0.001	< 0.0001
	LDL	-0.017	0.001	< 0.0001
	PCRln	-0.775	< 0.001	< 0.0001
	Age	-0.056	0.001	< 0.0001
	Hgb	1.049	< 0.001	< 0.0001
	BUN	-0.043	0.007	< 0.0001
	HbA1c	-0.069	< 0.001	< 0.0001
Event Process	Male	0.339	< 0.001	< 0.0001
	Diabetes	0.304	< 0.001	< 0.0001
	Hypertension	0.244	< 0.001	< 0.0001
	Assoct	-0.153	0.004	< 0.0001
	log(xi.1)	-5.964	< 0.001	
	log(xi.2)	-6.639	< 0.001	
	log(xi.3)	-21.137	< 0.001	
	log(xi.4)	-6.201	< 0.001	
	log(xi.5)	-4.929	< 0.001	
	log(xi.6)	-5.250	< 0.001	
	log(xi.7)	-5.553	< 0.001	

Table 4. Joint model results for groups 3b to 5.



Figure 3. Stage 3b–5 CKD example: late dialysis with the joint model. (**A**) Displays days since observation starts on the horizontal axis (0 to over 3000 days) and the probability of not undergoing dialysis on the vertical axis (0 to 1.0). The solid blue line represents the estimated probability of non-dialysis for case number 207, decreasing over time. Dashed black lines show the 95% confidence interval, widening over time to indicate increasing uncertainty in long-term forecasts. (**B–E**) Mark follow-up time at the top, with the solid green line for predicted GFR changes and dashed black lines for the 95% confidence interval of non-dialysis probability, indicating greater uncertainty over time. Asterisks (*) denote actual observed GFR values.



Figure 4. Stage 3b–5 CKD example: early dialysis with the joint model. (**A**) Displays days since observation starts on the horizontal axis (0 to over 3000 days) and the probability of not undergoing dialysis on the vertical axis (0 to 1.0). The solid blue line represents the estimated probability of non-dialysis for case number 96, decreasing over time. Dashed black lines show the 95% confidence interval, widening over time to indicate increasing uncertainty in long-term forecasts. (**B–E**) Mark follow-up time at the top, with the solid green line for predicted GFR changes and dashed black lines for the 95% confidence interval of non-dialysis probability, indicating greater uncertainty over time. Asterisks (*) denote actual observed GFR values.

Figure 4 contrasts sharply with Subject 96, which exemplifies an early dialysis onset within the same advanced stages of CKD. Here, the survival curve shows a rapid decline, with the blue solid line steeply dropping, indicating an urgent need for dialysis. The subplots, representing follow-up times (329, 420, 504, and 588 days), consistently display a steeper descent in non-dialysis probability, emphasizing the accelerated progression toward end-stage renal disease.

The juxtaposition of Figures 3 and 4 highlights the predictive versatility of the joint model across a spectrum of CKD severities. While both subjects fall under the same clinical classification of advanced CKD (stages 3b–5), the model discerns a significant difference in the timing of dialysis initiation. Subject 207's data project a more protracted course with a delayed approach to dialysis, whereas Subject 96's data signal a rapidly approaching need for dialytic intervention. This illustrates the model's robust capacity to individualize assessment.

6. Discussion

This retrospective cohort study focused on CKD patients engaged in disease management programs in Taiwan, aiming to delay the time to renal replacement therapy (RRT) through early detection of risk factors and providing healthcare education. Disease progression assessments can help clinicians plan for future situations, such as modifying a patient's CKD education, arranging vascular access creation, preparing for transplantation, or referring to hospice care [21]. Employing joint models, the research emphasized personalized CKD to hemodialysis progression prediction, illustrating the diversity of CKD patients' characteristics over time. The model's capability to differentiate the time to dialysis onset for patients within the same stage of CKD demonstrates its potential to inform individualized treatment decisions and underscores the need for a nuanced approach to patient care in chronic kidney disease management. Therefore, this joint model provides a nuanced understanding of the variables influencing renal function in patients with CKD. The significant parameters identified in both processes of the model offer valuable insights for clinicians in tailoring management strategies and in understanding the multifactorial nature of renal failure progression in this patient population and suggest its clinical value in guiding timely and appropriate therapeutic strategies for patients with varying prognoses within the advanced stages of CKD. Such personalized predictions are pivotal in optimizing patient outcomes and resource allocation in the management of CKD.

This study also highlighted that eGFR can be viewed as an endogenous factor influenced by innate biochemical markers, enhancing individual uncertainty management. This research's methodological framework focuses on the innovative use of joint models. These models are designed to clarify the correlation between eGFR temporal fluctuations and the risk of renal function deterioration. They are particularly suited for this task due to their ability to simultaneously analyze longitudinal data and time-to-event data. This dual approach provides a thorough understanding of eGFR dynamics in relation to renal failure risk. Furthermore, the utilization of risk prediction models can optimize healthcare resource allocation, targeting interventions toward high-risk CKD patients, potentially enhancing clinical outcomes and healthcare efficiency. This methodology offers a valuable tool for physicians and care workers to assess, intervene, and treat CKD progression promptly.

This understanding necessitates effective predictive tools, yet current methods face significant limitations. The prediction of CKD progression and renal failure constitutes a significant clinical challenge. Current predictive methods, including machine learning models [22,23] and traditional statistical approaches [24–26], though efficacious in certain aspects, do not adequately capture the dynamic and individual variability of renal function over time. These approaches frequently overlook comprehensive consideration of time-varying biomarker data and patient survival time data. Recognizing this gap, our study seeks to harness the power of joint models for more precise CKD outcomes. Initially developed for addressing issues in HIV research, this method has gradually been applied in other clinical research areas, including cancer, cardiovascular diseases, and kidney transplantation [16].

Enter the joint model, an innovative analytical tool that integrates longitudinal biomarker data, such as eGFR, with time-to-event data, such as the initiation of renal dialysis. While promising, the full potential of joint models in CKD assessment remains largely untapped. A key advantage of joint models is their intuitive interpretation of random effects in longitudinal submodels, especially when these models employ simple structures of random intercepts and slopes. This makes the models more interpretable and applicable statistically [27]. This framework leverages the joint model's unique ability to integrate multifaceted data. This methodology aims to surpass the constraints inherent in current predictive models, offering a more personalized and dynamic assessment of renal failure risk by incorporating patient-specific longitudinal data, including variations in eGFR and other pertinent clinical indicators. Our model endeavors to provide a nuanced understanding of CKD progression. This individualized approach is anticipated to significantly elevate the accuracy of risk assessments, thereby facilitating more effective and tailored patient management strategies in the realm of precision medicine [28].

This study found that serum creatinine, eGFR, blood pressure, urinary protein excretion, age, and gender have significant predictive abilities for the deterioration of renal function in both stages of CKD. These findings provide a comprehensive understanding of the factors influencing renal failure in stage 3–5 CKD patients, suggesting the potential for personalized treatment strategies tailored to individual risk profiles. The results of this study are comparable to previous reports [8,12,21,29,30]. Interestingly, age and sex are associated with CKD progression, regardless of whether in the early stage (3a) or in the pre-ESRD stage (3b–5). This study found that males were associated with a faster decline in eGFR than females. These results are compatible with many other studies [31–33]. This study also found that females are associated with a slower decline of eGFR and higher survival in patients or kidneys [31,34]. On the other hand, age also plays an important role, with older patients showing a more rapid decline, thus emphasizing early intervention's importance in younger patients. In a study of 129,486 adults in Canada with CKD stages 3a to 4 and followed for 6 years, the results found that when age increased, the likelihood of

CKD regression or death was greater than the likelihood of CKD progression or receiving dialysis therapy [35].

Strengths and Limitations

Our model's strength lies in its integration of various time-variant factors based on simultaneously analyzing longitudinal data and time-to-event data. These include eGFR changes, demographic details, and clinical parameters like blood pressure, blood glucose levels, lipid profiles, serum creatinine, and patient histories of hypertension, diabetes, or cardiovascular diseases. In terms of longitudinal data analysis, the mixed-effect model effectively tracks the changes in the same subject at different time points, such as eGFR and other clinical indicators, which helps in a more accurate understanding of the dynamic progression of CKD. However, a challenge of this mixed-effect model lies in potential issues with missing data and the complexity of the analysis. As for assembling with survival analysis, our model improves the challenge and is particularly suited for handling censored data, taking into consideration the duration of risk time, which is crucial for a deep understanding of time-to-event data. This integration enables a more accurate and dynamic interpretation of renal health trajectories. As a result, this approach is expected to substantially improve upon conventional models. It captures the complex and individualized patterns of CKD progression. The current model is applicable for prognostic assessment in clinical settings with patients' demographic characteristics similar to those in our study. Once the model is validated precisely, the web applications can be developed using the Shiny package for R.

Nonetheless, this study has certain limitations. It is a retrospective cohort study with a relatively small sample size, posing challenges in expanding the capability for model robusticity. To ensure unbiased results, it is imperative to conduct additional studies analyzing clinical pathological records from diverse hospitals, thereby increasing the sample size and improving the model's performance in predicting progression from early- and advanced-stage CKD. Additionally, the study's cohort had a mean age of 80 years, potentially limiting the generalizability of the findings to younger patients, which remains a concern when applying the model for earlier assessment of the patient's lifespan. Furthermore, future research needs to tackle the complexities inherent in joint models. This includes considering the handling of missing data and the potential influence of measurement errors, both of which are crucial for the model's robustness and reliability. Finally, model validation (e.g., cross-validation) involves assessing its performance and ensuring its reliability. This can be distinguished by creating a concordance index for joint models. Unfortunately, due to the relatively small sample size, dividing our current dataset into smaller subsets is not appropriate. In future research, the use of external samples for validation will be considered.

7. Conclusions

It is essential to take into account the diversities and trajectories of pathological indicators in the progression of pathology. In this study, the joint model, incorporating individual baseline characteristics and their variations over time in the progression, improves the conventional assessment models by employing simultaneously both the longitudinal analysis and survival analysis. This leads to a more precise assessment of renal failure risk, tailored to each patient's unique clinical profile. This model utilized a time-dependent endogenous factor (i.e., eGFR) and covariates, which is well-suited for the personalized assessment of trajectories at each follow-up measurement during clinic visits. Serum creatinine, blood pressure, and urinary protein excretion were identified as valuable factors for predicting the progression of CKD patients in the model. The approach holds promise for risk assessment in CKD progression, enabling healthcare professionals to identify individuals who could benefit from early intervention, such as timely referral for transplantation or initiation of dialysis. Author Contributions: Conceptualization, C.-M.L. (Chen-Mao Liao) and C.-M.L. (Chih-Ming Lin); Methodology, C.-M.L. (Chen-Mao Liao) and C.-M.L. (Chih-Ming Lin); Validation, C.-M.L. (Chen-Mao Liao) and Y.-W.K.; Formal analysis, C.-M.L. (Chen-Mao Liao); Investigation, Y.-P.C.; Resources, Y.-P.C.; Data curation, Y.-P.C.; Writing—original draft, Y.-W.K. and C.-M.L. (Chih-Ming Lin); Writing—review & editing, Y.-P.C. and C.-M.L. (Chih-Ming Lin); Visualization, Y.-W.K.; Project administration, C.-M.L. (Chih-Ming Lin); Funding acquisition, C.-M.L. (Chih-Ming Lin). All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Science and Technology Council grant number [110-2221-E-130-010].

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Taipei Veterans General Hospital (protocol code 2020-01-024BC).

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated and analyzed during the current study are not publicly available due to privacy/ethical restrictions but are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

- 1. Kovesdy, C.P. Epidemiology of chronic kidney disease: An update. Kidney Int. Suppl. 2022, 12, 7–11. [CrossRef]
- Jager, K.J.; Kovesdy, C.; Langham, R.; Rosenberg, M.; Jha, V.; Zoccali, C. A single number for advocacy and communication— Worldwide more than 850 million individuals have kidney diseases. *Kidney Int.* 2019, *96*, 1048–1050. [CrossRef] [PubMed]
- Bikbov, B.; Purcell, C.A.; Levey, A.S.; Smith, M.; Abdoli, A.; Abebe, M.; Owolabi, M.O. 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. [CrossRef] [PubMed]
- 4. Foreman, K.J.; Marquez, N.; Dolgert, A.; Fukutaki, K.; Fullman, N.; McGaughey, M.; Pletcher, M.A.; Smith, A.E.; Tang, K.; Yuan, C.W.; et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: Reference and alternative scenarios for 2016–40 for 195 countries and territories. *Lancet* **2018**, *392*, 2052–2090. [CrossRef] [PubMed]
- Saran, R.; Robinson, B.; Abbott, K.C.; Bragg-Gresham, J.; Chen, X.; Gipson, D.; Gu, H.; Hirth, R.A.; Hutton, D.; Jin, Y.; et al. US renal data system 2019 annual data report: Epidemiology of kidney disease in the United States. *Am. J. Kidney Dis.* 2020, 75, A6–A7. [CrossRef] [PubMed]
- 6. McCullough, K.P.; Morgenstern, H.; Saran, R.; Herman, W.H.; Robinson, B.M. Projecting ESRD incidence and prevalence in the United States through 2030. *J. Am. Soc. Nephrol.* **2019**, *30*, 127–135. [CrossRef] [PubMed]
- 7. Prasad, B.; Osman, M.; Jafari, M.; Gordon, L.; Tangri, N.; Ferguson, T.W.; Jin, S.; Kappel, J.; Kozakewycz, D. Kidney failure risk equation and cost of care in patients with chronic kidney disease. *Clin. J. Am. Soc. Nephrol.* **2022**, *17*, 17–26. [CrossRef]
- 8. Kuo, H.W.; Tsai, S.S.; Tiao, M.M.; Yang, C.Y. Epidemiological features of CKD in Taiwan. *Am. J. Kidney Dis.* 2007, 49, 46–55. [CrossRef]
- 9. Lee, C.C.; Hsu, C.C.; Lin, M.H.; Chen, K.H.; Wu, I.W. Hospitalization in patients with dialysis in Taiwan: A nationwide population-based observational study. *J. Formos. Med. Assoc.* **2022**, 121, S39–S46. [CrossRef]
- Levey, A.S.; Coresh, J.; Bolton, K.; Culleton, B.; Harvey, K.S.; Ikizler, T.A.; Johnson, C.A.; Kausz, A.; Kimmel, P.L.; Kusek, J.; et al. K/DOQI clinical practice guidelines for chronic kidney disease: Evaluation, classification, and stratification. *Am. J. Kidney Dis.* 2002, *39*, S1–S266.
- 11. Chittora, P.; Chaurasia, S.; Chakrabarti, P.; Culleton, B.; Harvey, K.S.; Ikizler, T.A.; Johnson, C.A.; Kausz, A.; Kimmel, P.L.; Kusek, J.; et al. Prediction of chronic kidney disease—A machine learning perspective. *IEEE Access* **2021**, *9*, 17312–17334. [CrossRef]
- 12. Tangri, N.; Stevens, L.A.; Griffith, J.; Tighiouart, H.; Djurdjev, O.; Naimark, D.; Levin, A.; Levey, A.S. A predictive model for progression of chronic kidney disease to kidney failure. *JAMA* **2011**, *305*, 1553–1559. [CrossRef] [PubMed]
- 13. Zhang, L.; Long, J.; Jiang, W.; Shi, Y.; He, X.; Zhou, Z.; Li, Y.; Yeung, R.O.; Wang, J.; Matsushita, K.; et al. Trends in chronic kidney disease in China. *N. Engl. J. Med.* **2016**, *375*, 905–906. [CrossRef]
- 14. Tangri, N.; Inker, L.A.; Hiebert, B.; Wong, J.; Naimark, D.; Kent, D.; Levey, A.S. A dynamic predictive model for progression of CKD. *Am. J. Kidney Dis.* **2017**, *69*, 514–520. [CrossRef]
- 15. Chang, Y.P.; Liao, C.M.; Wang, L.H.; Hu, H.H.; Lin, C.M. Static and dynamic prediction of chronic renal disease progression using longitudinal clinical data from Taiwan's national prevention programs. *J. Clin. Med.* **2021**, *10*, 3085. [CrossRef]
- 16. Asar, Ö.; Ritchie, J.; Kalra, P.A.; Diggle, P.J. Joint modelling of repeated measurement and time-to-event data: An introductory tutorial. *Int. J. Epidemiol.* **2015**, *44*, 334–344. [CrossRef] [PubMed]
- 17. Levey, A.S.; Bosch, J.P.; Lewis, J.B.; Greene, T.; Rogers, N.; Roth, D. A more accurate method to estimate glomerular filtration rate from serum creatinine: A new prediction equation. *Ann. Intern. Med.* **1999**, *130*, 461–470. [CrossRef]

- Schwandt, A.; Denkinger, M.; Fasching, P.; Pfeifer, M.; Wagner, C.; Weiland, J.; Zeyfang, A.; Holl, R.W. Comparison of MDRD, CKD-EPI, and Cockcroft-Gault equation in relation to measured glomerular filtration rate among a large cohort with diabetes. J. Diabetes Complicat. 2017, 31, 1376–1383. [CrossRef] [PubMed]
- 19. Snyder, S.; Pendergraph, B. Detection and evaluation of chronic kidney disease. Am. Fam. Physician 2005, 72, 1723–1732.
- 20. Rubin, D.B.; Schenker, N. Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *J. Am. Stat. Assoc.* **1986**, *81*, 366–374. [CrossRef]
- Ramspek, C.L.; Evans, M.; Wanner, C.; Drechsler, C.; Chesnaye, N.C.; Szymczak, M.; Krajewska, M.; Torino, C.; Porto, G.; Hayward, S.; et al. Kidney failure prediction models: A comprehensive external validation study in patients with advanced CKD. J. Am. Soc. Nephrol. 2021, 32, 1174. [CrossRef]
- 22. Bou-Matar, R.; Dell, K.M.; Bobrowski, A. Machine learning models to predict post-dialysis blood pressure in children and young adults on maintenance hemodialysis. *Sci. Rep.* 2023, *13*, 19105. [CrossRef]
- 23. Rankin, S.; Han, L.; Scherzer, R.; Tenney, S.; Keating, M.; Genberg, K.; Rahn, M.; Wilkins, K.; Shlipak, M.; Estrella, M. A Machine learning model for predicting mortality within 90 days of dialysis initiation. *Kidney360* **2022**, *3*, 1556–1565. [CrossRef]
- 24. Liu, L.; Pang, J.; Xu, J.; Liu, L.N.; Liao, M.Y.; Huang, Q.X.; Li, Y.L. Impact of initial dialysis modality on the survival of patients with ESRD: A propensity-score-matched study. *BMC Nephrol.* **2023**, *24*, 313. [CrossRef]
- 25. Pitanupong, J.; Chang, A. Relationship of serum albumin at initiation of incident peritoneal dialysis with cardiovascular and overall survival. *Clin. Nephrol.* **2023**, *99*, 265–273. [CrossRef]
- Ma, X.Y.; Sheng, Y.P.; Yang, X.M.; Wang, N.; Zhang, H.R.; Xu, H.P.; Sun, F.Y. Characteristics and prognostic outcome factors between young and elderly peritoneal dialysis patients: A prospective cohort study. *Ann. Palliat. Med.* 2022, 11, 2952–2960. [CrossRef]
- 27. Dong, J.; Cao, J.; Gill, J.; Miles, C.; Plumb, T. Functional joint models for chronic kidney disease in kidney transplant recipients. *Stat. Methods Med. Res.* **2021**, *30*, 1932–1943. [CrossRef] [PubMed]
- Zacharias, H.U.; Altenbuchinger, M.; Schultheiss, U.T.; Raffler, J.; Kotsis, F.; Ghasemi, S.; Ali, I.; Kollerits, B.; Metzger, M.; Steinbrenner, I.; et al. A predictive model for progression of CKD to kidney failure based on routine laboratory tests. *Am. J. Kidney Dis.* 2022, *79*, 217–230. [CrossRef] [PubMed]
- Hannan, M.; Ansari, S.; Meza, N.; Anderson, A.H.; Srivastava, A.; Waikar, S.; Charleston, J.; Weir, M.R.; Taliercio, J.; Horwitz, E.; et al. Risk factors for CKD progression: Overview of findings from the CRIC study. *Clin. J. Am. Soc. Nephrol.* 2021, 16, 648. [CrossRef] [PubMed]
- 30. Segal, Z.; Kalifa, D.; Radinsky, K.; Ehrenberg, B.; Elad, G.; Maor, G.; Lewis, M.; Tibi, M.; Korn, L.; Koren, G.; et al. Machine learning algorithm for early detection of end-stage renal disease. *BMC Nephrol.* **2020**, *21*, 518. [CrossRef] [PubMed]
- Eriksen, B.O.; Ingebretsen, O.C. The progression of chronic kidney disease: A 10-year population-based study of the effects of gender and age. *Kidney Int.* 2006, 69, 375–382. [CrossRef]
- 32. Ricardo, A.C.; Yang, W.; Sha, D.; Appel, L.J.; Chen, J.; Krousel-Wood, M.; Manoharan, A.; Steigerwalt, S.; Wright, J.; Rahman, M.; et al. Sex-related disparities in CKD progression. *J. Am. Soc. Nephrol.* **2019**, *30*, 137. [CrossRef]
- Minutolo, R.; Gabbai, F.B.; Chiodini, P.; Provenzano, M.; Borrelli, S.; Garofalo, C.; Bellizzi, V.; Russo, D.; Conte, G.; De Nicola, L. Sex differences in the progression of CKD among older patients: Pooled analysis of 4 cohort studies. *Am. J. Kidney Dis.* 2020, 75, 30–38. [CrossRef]
- 34. Swartling, O.; Rydell, H.; Stendahl, M.; Segelmark, M.; Lagerros, Y.T.; Evans, M. CKD progression and mortality among men and women: A nationwide study in Sweden. *Am. J. Kidney Dis.* **2021**, *78*, 190–199. [CrossRef]
- Liu, P.; Quinn, R.R.; Lam, N.N.; Al-Wahsh, H.; Sood, M.M.; Tangri, N.; Tonelli, M.; Ravani, P. Progression and regression of chronic kidney disease by age among adults in a population-based cohort in Alberta, Canada. *JAMA Netw. Open* 2021, 4, e2112828. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.