



# Statistical Methods for Cohort Studies of CKD: Survival Analysis in the Setting of Competing Risks

Jesse Yen-Chih Hsu,<sup>\*†</sup> Jason A. Roy,<sup>\*†</sup> Dawei Xie,<sup>\*†</sup> Wei Yang,<sup>\*†</sup> Haochang Shou,<sup>\*†</sup> Amanda Hyre Anderson,<sup>\*†</sup> J. Richard Landis,<sup>\*†</sup> Christopher Jepson,<sup>\*†</sup> Myles Wolf,<sup>‡</sup> Tamara Isakova,<sup>§||</sup> Mahboob Rahman,<sup>\*†\*\*††</sup> and Harold I. Feldman,<sup>\*†</sup> and on behalf of the Chronic Renal Insufficiency Cohort (CRIC) Study Investigators

## Abstract

Survival analysis is commonly used to evaluate factors associated with time to an event of interest (e.g., ESRD, cardiovascular disease, and mortality) among CKD populations. Time to the event of interest is typically observed only for some participants. Other participants have their event time censored because of the end of the study, death, withdrawal from the study, or some other competing event. Classic survival analysis methods, such as Cox proportional hazards regression, rely on the assumption that any censoring is independent of the event of interest. However, in most clinical settings, such as in CKD populations, this assumption is unlikely to be true. For example, participants whose follow-up time is censored because of health-related death likely would have had a shorter time to ESRD, had they not died. These types of competing events that cause dependent censoring are referred to as competing risks. Here, we first describe common circumstances in clinical renal research where competing risks operate and then review statistical approaches for dealing with competing risks. We compare two of the most popular analytical methods used in settings of competing risks: cause-specific hazards models and the Fine and Gray approach (subdistribution hazards models). We also discuss practical recommendations for analysis and interpretation of survival data that incorporate competing risks. To demonstrate each of the analytical tools, we use a study of fibroblast growth factor 23 and risks of mortality and ESRD in participants with CKD from the Chronic Renal Insufficiency Cohort Study.

*Clin J Am Soc Nephrol* 12: 1181–1189, 2017. doi: <https://doi.org/10.2215/CJN.10301016>

## Introduction

Survival analysis refers to a class of statistical analytical methods that are used when the outcome variable is the time until a specified event of interest occurs. In clinical research, the event of interest is typically a clinical outcome, such as ESRD. Survival analysis can be a powerful tool to help us to understand the associations between risk factors and clinical outcomes, and also to predict an individual's risk of developing a clinical outcome. Like many other statistical techniques, however, care must be taken to ensure that the specific method selected for the analysis is appropriate to the situation; otherwise, erroneous conclusions may be drawn.

This paper examines analytical approaches in clinical settings where there are competing risks. Competing risks exist when an individual can experience other types of events that prevent the occurrence of the event of interest, thereby preventing us from knowing when that individual would have developed the event of interest had the competing event not have happened (1). For example, death precludes us from observing CKD outcomes such as ESRD (2,3), or cardiovascular disease (CVD) outcomes such as coronary artery disease or congestive heart failure (4,5).

Without additional adjustments, certain commonly used survival analysis methods, such as the Kaplan–Meier survival function and Cox proportional hazards regression, can produce misleading conclusions when

competing risks are present, which causes informative censoring or incorrect event probabilities. Consequently, various approaches to dealing with this problem have been developed. The goal of this paper is four-fold. First, we explain the concept of competing risks and outline the implications of competing risks for standard survival analysis methods, such as the Kaplan–Meier estimator and Cox proportional hazards models. Second, we provide a brief explanation of two main types of survival models designed for use in the presence of competing risks: cause-specific hazards models and subdistribution hazards models. Third, we discuss key considerations in planning statistical analyses in the presence of competing risks, on the basis of the specific scientific questions the analysis is intended to address. Fourth, we use a motivating example in CKD to illustrate practical and analytical considerations for analysis and interpretation of survival data that incorporate competing risks.

## Motivating Example: Is Fibroblast Growth Factor 23 Associated with Time to ESRD?

The motivating example (2) is an analysis of data from 3879 participants enrolled between June of 2003 and September of 2008 in the Chronic Renal Insufficiency Cohort (CRIC) Study, a prospective study of individuals with stage 2–4 CKD (6–8). The goal of the motivating study was to evaluate fibroblast growth factor 23 (FGF-23) as a risk factor for developing ESRD

<sup>\*</sup>Department of Biostatistics and Epidemiology and <sup>†</sup>Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; <sup>‡</sup>Department of Medicine, School of Medicine, Duke University, Durham, North Carolina; <sup>§</sup>Division of Nephrology and Hypertension, Department of Medicine, Northwestern University, Chicago, Illinois; <sup>||</sup>Center for Translational Metabolism and Health, Northwestern University Feinberg School of Medicine, Chicago, Illinois; and <sup>¶</sup>Division of Nephrology and Hypertension, Case Western Reserve University, Cleveland, Ohio; <sup>\*\*</sup>University Hospitals Cleveland Medical Center, Cleveland, Ohio; and <sup>††</sup>Louis Stokes Cleveland Veterans Affairs Medical Center, Cleveland, Ohio

## Correspondence:

Dr. Jesse Y. Hsu, Department of Biostatistics and Epidemiology and Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, 618 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104. Email: [hsu9@upenn.edu](mailto:hsu9@upenn.edu)

in participants with CKD (2). In this example, death can be viewed as an event competing with ESRD. During a median follow-up time of 3.5 years, 410 study participants developed ESRD and 246 participants died without developing ESRD. The remaining 3223 participants experienced neither ESRD nor death while they were followed in the CRIC study until the end of the analysis. Median FGF-23 levels were 146 (interquartile range, 96–239) RU/ml. At baseline, the mean age was 57 years and mean eGFR was 43 ml/min per 1.73 m<sup>2</sup>. In the cohort, 45% of participants were women, 45% were non-Hispanic white, 42% were non-Hispanic black, 13% were Hispanic, and 48% had diabetes. For additional details about this analysis, see Isakova *et al.* (2).

### Introduction to Classic Survival Analysis for a Single Event

To facilitate understanding the challenges of, and approaches to, analyzing data with competing risks, we first provide a brief overview of survival analysis. As described above, in survival analysis, the outcome is the time until a specified event of interest occurs. For each individual in the analysis, the time from start of follow-up until the time the event occurs is referred to as that individual's survival time.

### Independent Censoring and Competing Events

In an ideal situation in which all subjects in the study were followed until they experienced the event of interest, we would have an observed time to event for each person and could model it using standard methods (*e.g.*, regression modeling). In practice, however, this rarely occurs; some participants never experience the event during follow-up. When a subject has follow-up end before experiencing the event of interest, the subject is censored. Censoring can happen for any of a number of reasons, including death, loss to follow-up, and end of study.

The various methods of survival analysis are designed to handle data sets that include censored cases. Two commonly used methods, the Kaplan–Meier estimator and Cox proportional hazards models, are described below.

#### Kaplan–Meier Estimator

The Kaplan–Meier estimator is the most frequently used tool to estimate the survival function (9). The Kaplan–Meier estimator is a step-down function, with each step at each time an event of interest occurs. The height of a step at a given time is the proportion of subjects at risk (*i.e.*, free of the event of interest) just before the given time, who experience the event at that time. The log-rank test can be used to compare two or more Kaplan–Meier estimators on the basis of different values of a single covariate; for example, treated versus control groups, smokers versus nonsmokers, men versus women, *etc.*

#### Cox Proportional Hazards Models

Proportional hazards models, or Cox proportional hazards models (10), allow us to investigate the association between a set of covariates and the event of interest. The Cox proportional hazards model estimates the hazard function (also called hazard rate),  $h(t)$ , which is the

instantaneous rate of occurrence of the event of interest at survival time,  $t$ , for a subject who has survived event free up until time  $t$ . The term “instantaneous rate” refers to the event rate in a short period of time. For example,  $h(2.5)$  is the risk of developing ESRD just after year 2.5 for subjects who have not yet developed ESRD. In a Cox proportional hazards analysis, the effect of each covariate is found by estimating the hazard rates among subjects differing by one unit on the given covariate and taking the ratio of the two rates; this ratio is familiarly known as the hazard ratio (see Supplemental Table 1 for technical details on the hazard function and other functions commonly used in survival analysis).

An important assumption made by the Cox proportional hazards model is the proportional hazards assumption. This assumption states that the hazard ratio for a given covariate remains constant over time. Violation of proportionality in Cox proportional hazards models may lead to misleading conclusions; thus, it is useful to check proportionality (*e.g.*, Schoenfeld residuals or time-dependent coefficients—an interaction between a function of time and predictors) in practice. In the presence of violation of proportionality in Cox proportional hazards model, adding the time-dependent coefficients would allow us to estimate the hazard ratios over time.

One concept that is fundamental to understanding how the Cox proportional hazards model (and survival analysis generally) works is that of the “risk set.” The risk set is the set of subjects who, at any given point in time, have not yet experienced the event of interest and could potentially be observed to experience that event in the future. At the beginning of the study, the risk set includes all subjects; over time, subjects experiencing the event of interest, and censored subjects, are excluded from the risk set. At each point in time, the Cox proportional hazards model uses the data from all cases that are in the risk set at that particular time.

### Analytical Strategies and Key Assumptions in the Presence of a Competing Event

The methods of survival analysis described above assume that censoring occurs independently of the event of interest. A censoring event is said to be independent of the event of interest if the likelihood of occurrence of the two events are not correlated. For example, consider a study in which the event of interest is ESRD. If a participant in such a study dies in a car crash before developing ESRD, this obviously prevents us from knowing how long it would have taken that individual to reach ESRD, had they lived. However, there is no reason to assume that individuals who die in car crashes would, had they lived, have systematically different (shorter or longer) times to ESRD than other participants, so the two events are considered to be independent.

The assumption of independent censoring simplifies survival analysis because it allows us to treat those cases for whom survival times are not known (*i.e.*, censored cases) as being the same as cases for whom survival times are known. As a result, methods of survival analysis that rest on the assumption of independent censoring, such as the Kaplan–Meier estimator and the standard Cox

proportional hazards model, are widely used. However, there are many types of censoring events that may not be independent of the event of interest. For example, in the study of ESRD just described, if participants who have more advanced CKD (and are thus more likely to experience ESRD soon) are also more likely to have CVD, then CVD-related death is not an independent censoring event. In situations of this type, methods such as the Kaplan–Meier estimator and the standard Cox proportional hazards model can yield biased results, because they treat censored individuals as having the same risk of the event of interest as uncensored individuals. Unfortunately, the assumption of independent censoring is, in general, not testable.

In the example given above, CVD-related death is an example of a competing risk event, as it precludes any possibility of the subsequent occurrence of ESRD. Table 1 provides examples of differences in competing events and censoring events. Competing risk events are quite commonly not independent of the event of interest. Consequently, analytic techniques that have been developed for use in the setting of competing risks do not rest on the assumption of independent censoring. In this section, we review two such techniques, cause-specific hazards models and subdistribution hazards models, which yield interpretable measures of association between covariates and the event of interest, and of prediction of the event of interest in the presence of competing events.

**Cause-Specific Hazards Models**

When there are competing events, the Cox proportional hazards approach can still be used by treating the competing event as a censoring event. Prentice *et al.* (11) proposed a simple method for doing this. In this approach, separate models are performed to estimate the hazard rates for each specific cause of failure (*i.e.*, the event of interest and competing events). Because these models are for specific causes, they are referred to as cause-specific hazards models.

In practical terms, the primary impact of cause-specific hazards models is the interpretation of the hazard ratios. Because any competing event is treated like other independent censoring events in cause-specific hazards models, the cause-specific hazard ratio,  $h_{1,CS}(t)$ , is the instantaneous rate of occurrence of the event of interest at survival time,  $t$ , for a subject who has survived event free (both the event of interest and competing events) up until time  $t$ , and  $HR_{1,CS}$ , is interpreted as the hazard ratio for the event of interest comparing two values of a factor in “a hypothetical world,”

where the competing event is removed. In other words,  $HR_{1,CS}$  is what the hazard ratio for the event of interest would be if subjects could never experience the competing event. Similar to Cox proportional hazards models, cause-specific hazards models require the proportionality assumption such that the cause-specific hazard ratio is assumed to remain constant over time. Lunn and McNeil (12) proposed a joint stratified (by event types) Cox proportional hazards model with rearrangement of the data set to replace separate cause-specific hazards models, and many authors have implemented this approach in practice (13,14).

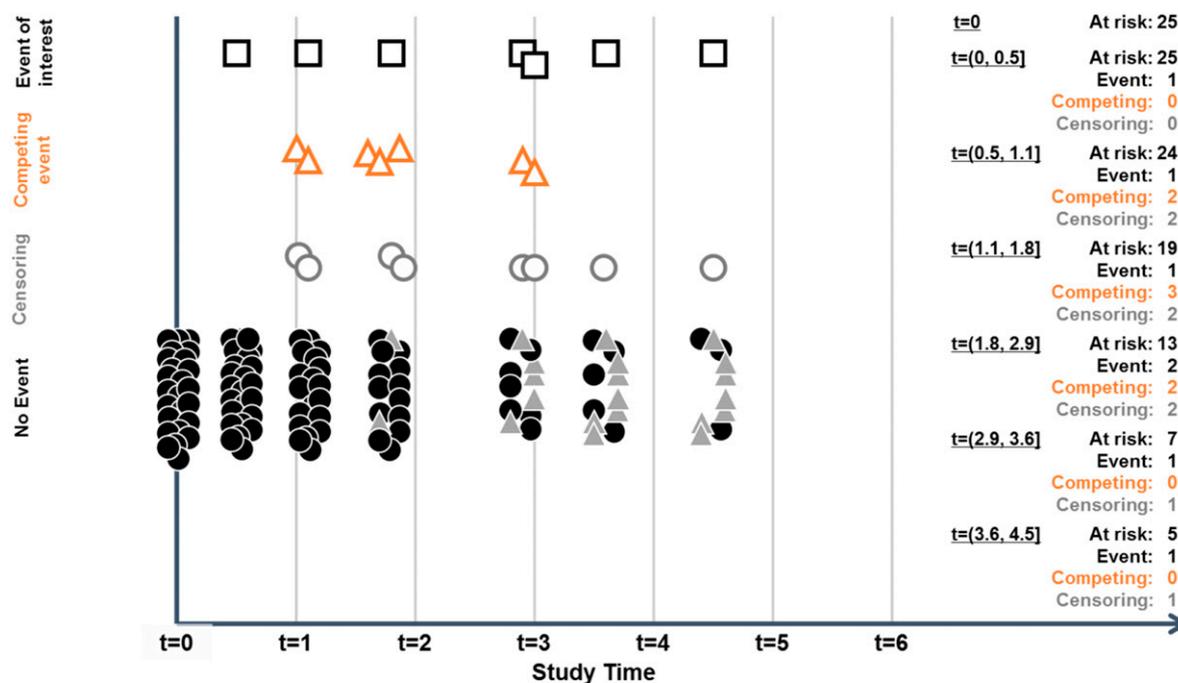
**Subdistribution Hazards Models**

Although the concept of cause-specific hazards models is straightforward, researchers often find the caveat “in a hypothetical world where the competing event is removed” in the interpretation of  $HR_{1,CS}$  opaque. Fine and Gray proposed an approach that removes this caveat (15). In technical terms, the primary difference between cause-specific hazards models and the Fine and Gray approach is that the Fine and Gray approach models hazards on the basis of the cumulative incidence function,  $F(t)$ , which is the risk of a subject experiencing the event of interest before survival time  $t$ ; hazards modeled in this way are referred to as subdistribution hazards (see Supplemental Table 1 for technical details on the cumulative incidence function). In practical terms, the primary difference between cause-specific hazards models and the Fine and Gray approach is in their operational differences: that is, how subjects who have experienced a competing event are treated in the risk set. In cause-specific hazards models, as we have seen previously, such subjects are removed from the risk set. In the Fine and Gray approach, such subjects are retained in the risk set after experiencing the competing event. Thus, the risk set includes subjects for whom the event of interest can potentially be observed, as well as subjects for whom the event of interest cannot be observed because of the previous occurrence of a competing event from the risk set.

As noted by other authors (15,16), the cause-specific hazard represents the instantaneous rate of occurrence of the event of interest at  $t$  for a subject who is still at risk before  $t$ , whereas the subdistribution hazard shows the instantaneous rate of occurrence of the event of interest at  $t$  for a subject who either is still at risk or has experienced the competing event before  $t$ . Following Lau *et al.* (17), Figure 1 depicts the difference in risk sets between cause-specific hazards models and subdistribution hazards models in a hypothetical example of 25 subjects. In Figure 1, after seven subjects experienced the competing event (triangles) at time 1.1 (two), time 1.8 (three) and time 2.9 (two), they stay in the risk set for the remaining time in subdistribution hazards, while they are removed from the risk set for the remaining time in cause-specific hazards. The difference in the risk sets between cause-specific and subdistribution hazards can be seen in the denominators in the table in Figure 1. Like cause-specific hazards models, subdistribution hazards models also require the proportionality assumption; *i.e.*, a subdistribution hazard ratio for any two values of a factor remains constant over time. Fine and Gray (15) also showed that a hazard ratio–like estimator,

**Table 1. Comparison of examples for censoring and competing events in CKD**

Censoring	Competing Event	Noncompeting Event
Noninformative	Car crash death	End of study
Informative	Cardiovascular death	Loss to follow-up (health status)



Events/Risk set	t = 0	t = 0.5	t = 1.1	t = 1.8	t = 2.9	t = 3.6	t = 4.5
Cause-specific	0/25	1/25	1/24	1/19	2/13	1/7	1/5
Sub-distribution	0/25	1/25	1/24	1/21	2/18	1/14	1/14

**Figure 1.** | In a hypothetical example of 25 subjects, the cause-specific hazard function for the event of interest at time  $t$  and the sub-distribution hazard function for the event of interest at time  $t$  are calculated using different risk sets (i.e., differences in their denominators). Solid circles: subjects without experiencing any events; squares: subjects experiencing the event of interest; empty (orange) triangles: subjects experiencing the competing event; solid (gray) triangle: subjects experiencing the competing event in earlier time points and stayed in the risk set; circles: subjects being censored.

the subdistribution hazard ratio, denoted as  $\text{SHR}_1$  and interpreted as the hazard ratio for the event of interest in a setting where subjects can experience competing events, can be derived from the models.

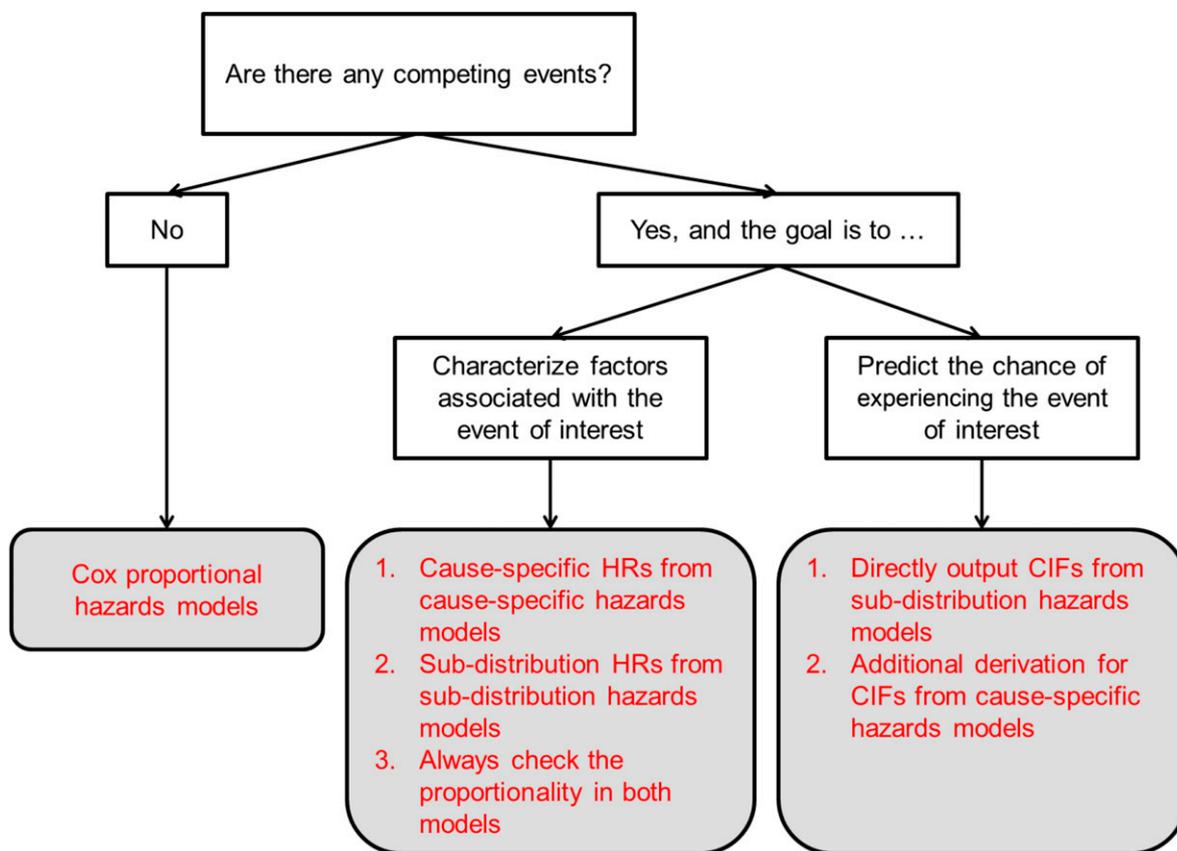
### Comparing Cause-Specific and Subdistribution Hazards Models

Although modern statistical packages allow us to perform analyses of cause-specific hazards models and sub-distribution hazard models easily, it is worth considering some analytical strategies on the basis of two main goals, association and prediction. In Figure 2, we show a diagram to guide analytical strategies on the basis of the goal of the analysis.

If the goal is to understand the association between a predictor and the event of interest, then the use of cause-specific hazards models, with interpretation of cause-specific hazard ratios, can be useful to inform the understanding of etiology, the relationship between the predictor and the event of interest assuming the competing events do not exist. The use of subdistribution hazards models, with interpretation of subdistribution hazard ratios, can estimate the relationship between the predictor and the event of interest while taking into account the competing events.

If, on the other hand, the goal is to predict the chance that a subject with certain characteristics will experience the event of interest within a given time period, sub-distribution hazards models provide a convenient and direct calculation of cumulative incidence functions. Although the cumulative incidence functions cannot be obtained directly from cause-specific hazards models, they can still be derived from integration of product of cause-specific hazard and overall survival functions (18).

In situations where researchers may be tempted to perform both cause-specific and subdistribution hazards models, it is recommended to pay additional attention to the proportionality assumptions. We have mentioned that proportionality assumptions are necessary in both cause-specific and subdistribution hazards models, however, as many authors have pointed out, the proportionality in cause-specific hazards models and the proportionality in subdistribution hazards models cannot be satisfied simultaneously without restrictions (19,20). To describe the potential impact of the competing events on the event of interest, a time-dependent coefficient, an interaction term between the exposure of interest and a function of survival time (e.g., logarithm of time or dichotomization at a certain time point) can be added to both models to allow hazard ratios to change over time. Supplemental



**Figure 2.** | A simple flow chart to illustrate analytical strategies on the basis of the goal of the analysis. CIF, cumulative incidence function; HR, hazard ratio.

Table 2 shows technical details of the relationship between cause-specific and subdistribution hazard ratios under different restrictions. Muñoz *et al.* (20) presented a more formal definition for readers who are interested in further details.

**A Motivating Example in the Presence of Competing and Independent Censoring Events**

In this section, we consider a binary primary predictor, FGF-23 > 180 RU/ml versus FGF-23 ≤ 180 RU/ml. Let ESRD be the event of interest and let loss to follow-up and end of the study be censoring events. Because death precludes us from observing ESRD, we treat death as a competing event in all analyses. Table 2 shows the distributions of death and ESRD from 3879 participants by subgroups of FGF-23.

Subgroup	Censoring	ESRD	Death	Total
FGF-23 ≤ 180 RU/ml	2181	135	95	2411
FGF-23 > 180 RU/ml	1042	275	151	1468
Total	3223	410	246	3879

There were 410 participants who had developed ESRD and 246 participants who had died during the study period; 3223 participants were either lost to follow-up or did not have any event before the end of the analysis.

**Association between FGF-23 and Time to ESRD**

To characterize the association between levels of FGF-23 and ESRD, we use hazard ratios from the cause-specific and subdistribution hazard models with and without a time-dependent coefficient for levels of FGF-23, which is defined as an interaction between levels of FGF-23 and logarithm of survival time, in Table 3. From the cause-specific hazards model (*i.e.*, relationship between levels of FGF-23 and ESRD when the competing event, death, would not occur), the significant time-dependent coefficient for levels of FGF-23 ( $P=0.002$ ) indicates that the cause-specific hazard ratio for ESRD,  $HR_{ESRD, CS}$ , changes over time. For example,  $HR_{ESRD, CS}$  is 5.90 (with a 95% confidence interval of 4.21 to 8.27) at year 1 and  $HR_{ESRD, CS}$  is 2.73 (with 95% confidence interval of 2.00 to 3.73) at year 4. From the subdistribution hazards model (*i.e.*, relationship between levels of FGF-23 and ESRD when taking into account the competing event, death), the significant time-dependent coefficient for levels of FGF-23 ( $P=0.002$ ) indicates that the subdistribution hazard ratio for ESRD,  $SHR_{ESRD}$ , will change over time. Figure 3 depicts  $HR_{ESRD, CS}$  and  $SHR_{ESRD}$  and their 95% confidence intervals from the cause-specific and subdistribution hazards models with a time-dependent

**Table 3. Analyses of cause-specific and subdistribution hazards models for characterizing levels of FGF-23 associated with development of ESRD from the motivating example**

ESRD	Cause-Specific Hazards Models		Subdistribution Hazards Models	
	Coefficient (P Value)	HR <sub>ESRD, CS</sub> (95% CI)	Coefficient (P Value)	SHR <sub>ESRD</sub> (95% CI)
<b>With time-dependent coefficient for FGF-23</b>				
FGF-23>180 RU/ml	1.78 (<0.001)	5.90 (4.21 to 8.27) at year 1	1.75 (<0.001)	5.76 (4.17 to 7.97) at year 1
Interaction of FGF-23>180 RU/ml and time	-0.56 (0.002)	4.02 (3.25 to 4.96) at year 2 3.21 (2.51 to 4.10) at year 3 2.73 (2.00 to 3.73) at year 4 2.41 (1.66 to 3.51) at year 5	-0.62 (0.002)	3.76 (3.05 to 4.65) at year 2 2.93 (2.30 to 3.75) at year 3 2.46 (1.81 to 3.34) at year 4 2.14 (1.45 to 3.08) at year 5
<b>Without time-dependent coefficient for FGF-23<sup>a</sup></b>				
FGF-23>180 RU/ml	1.39 (<0.001)	4.01 (3.26 to 4.93)	1.33 (<0.001)	3.77 (3.07 to 4.62)

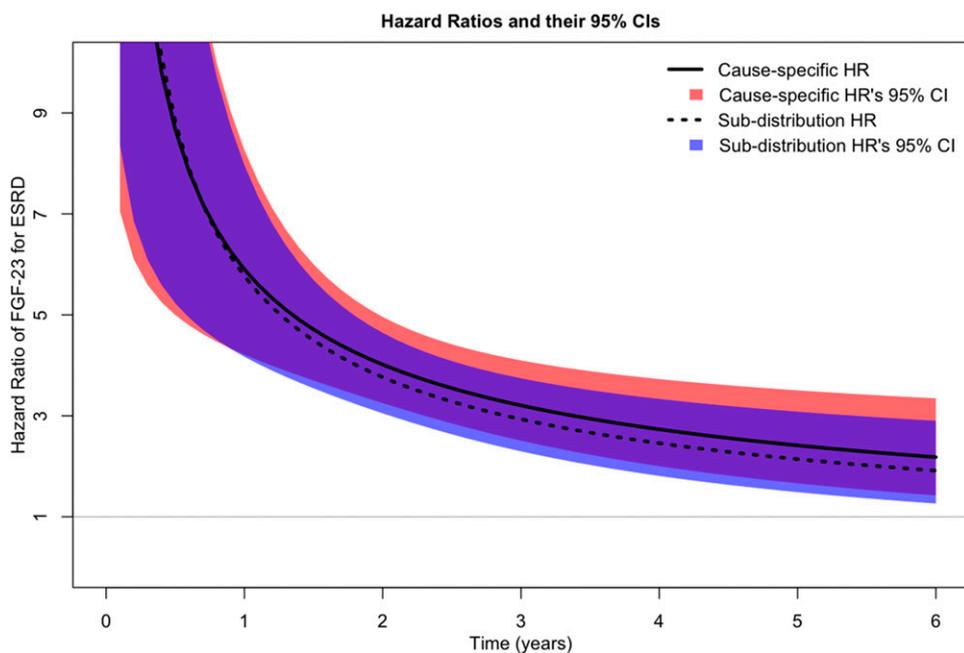
For the purpose of illustration, all models are unadjusted. Time-dependent FGF-23>180 RU/ml is defined as an interaction between FGF-23>180 RU/ml and logarithm of survival time. FGF-23, fibroblast growth factor 23; HR<sub>ESRD, CS</sub>, cause-specific hazard ratio for ESRD; 95% CI, 95% confidence interval; SHR<sub>ESRD</sub>, sub-distribution hazard ratio for ESRD.

<sup>a</sup>Models without time-dependent coefficient for levels of FGF-23 are presented here for the purpose of comparison. We do not need them in practice because of the presence of significant time-dependent coefficient for levels of FGF-23.

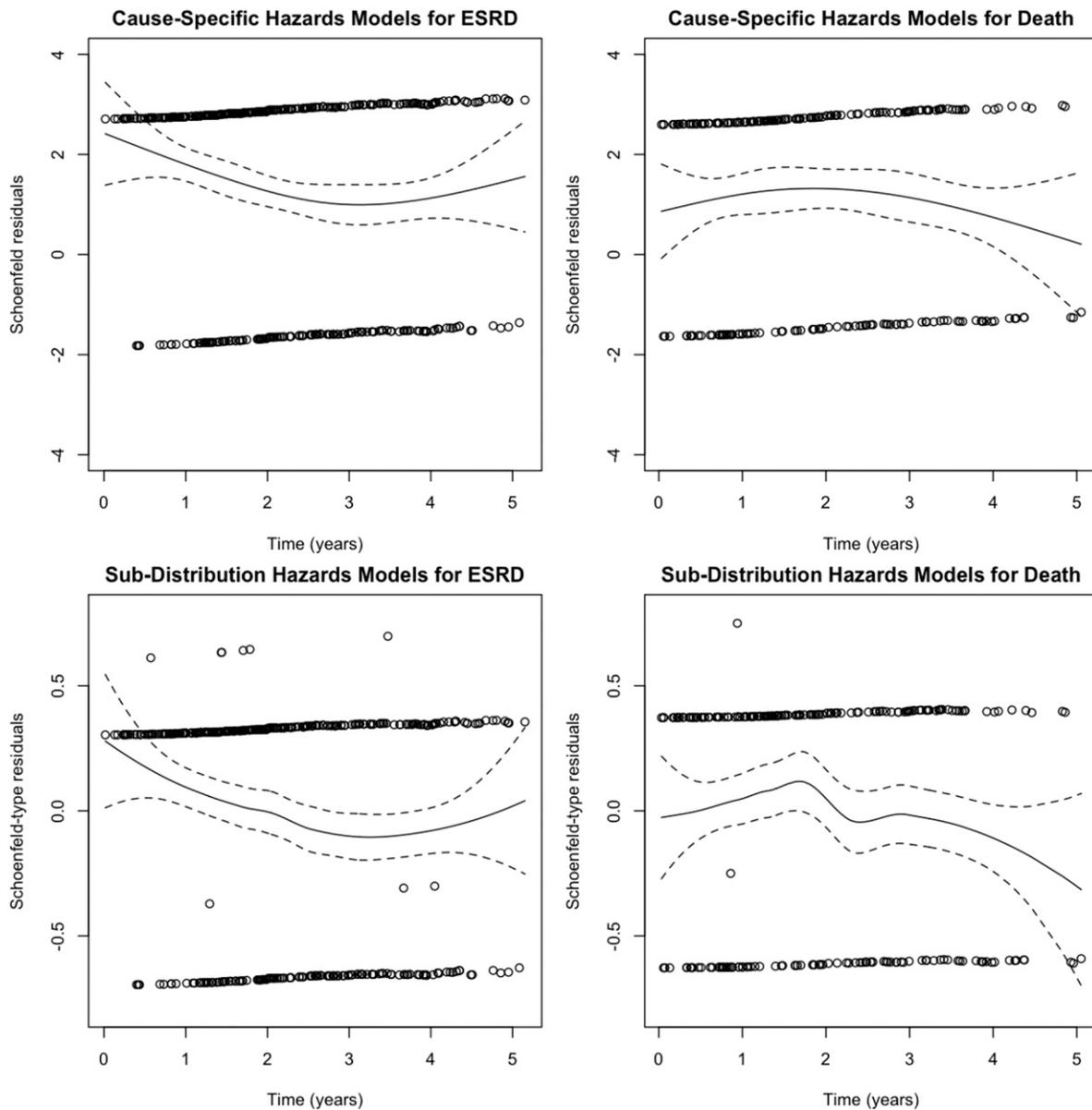
coefficient. Both HR<sub>ESRD, CS</sub> and SHR<sub>ESRD</sub> are higher in the early years and decrease over time. Although the majority of the confidence intervals for HR<sub>ESRD, CS</sub> and SHR<sub>ESRD</sub> are overlapped, HR<sub>ESRD, CS</sub> and SHR<sub>ESRD</sub> are similar within the

first year and HR<sub>ESRD, CS</sub> is slightly larger than SHR<sub>ESRD</sub> in later years.

Another way to check the proportionality assumption is to plot residuals against time. Figure 4 illustrates four



**Figure 3. | Estimated cause-specific hazard ratio (solid line) and subdistribution hazard ratio (dashed line) over time and their corresponding 95% confidence intervals (red area: cause-specific hazard ratio; blue area: subdistribution hazard ratio) from models including a time-dependent coefficient for levels of FGF-23 to allow hazard ratios to change over time. 95% CI, 95% confidence interval; FGF-23, fibroblast growth factor 23; HR, hazard ratio.**



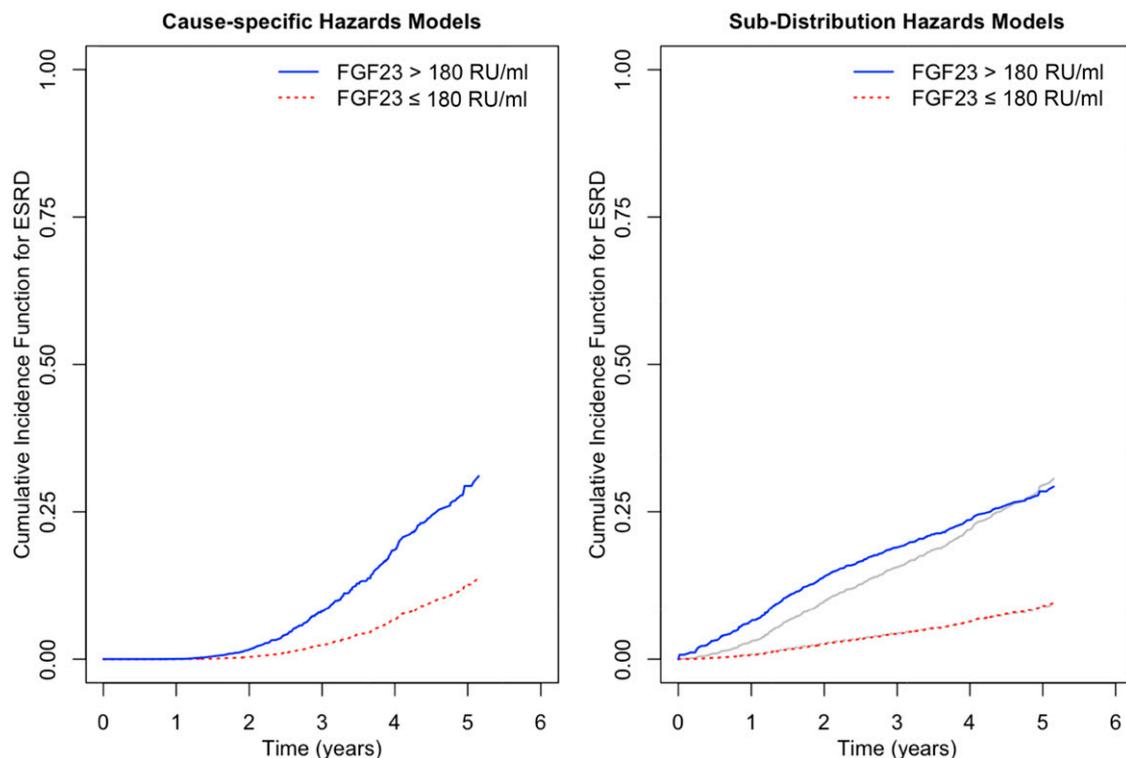
**Figure 4.** | Schoenfeld residuals plots for cause-specific hazards models and Schoenfeld-type residuals plots for subdistribution hazards models. Circles: Schoenfeld residuals; lines: smooth curve; dashed lines: confidence bands at two standard errors.

residuals plots: (1) two Schoenfeld residuals plots from cause-specific hazards models for ESRD and death (top); and (2) two Schoenfeld-type residuals plots from subdistribution hazards models for ESRD and death (bottom). All models do not include a time-dependent coefficient. Smooth curves in four residuals plots reveal evidence of violating the proportionality assumption. Thus, models with a time-dependent coefficient for levels of FGF-23 in Table 3 are more appropriate than those without the time-dependent coefficient.

**Prediction of Developing ESRD**

To predict the chance of developing ESRD for participants on the basis of levels of FGF-23, we will use cumulative incidence functions. Because subdistribution hazards models are built directly on cumulative incidence

functions, it is straightforward to obtain cumulative incidence functions once the models are fitted. Figure 5 presents the cumulative incidence functions for participants with FGF-23 > 180 RU/ml (solid blue line) and for participants with FGF-23 ≤ 180 RU/ml (dashed red line) on the basis of the cause-specific (left) and subdistribution (right) hazards models. Note that both approaches provide similar estimated cumulative incidence functions for ESRD (see Supplemental Figure 1 for SAS programming examples for calculating cumulative incidence functions on the basis of both models; for more details, see Cheng *et al.* [18] and So *et al.* [21]). By the end of year 5, about 29% of participants with FGF-23 > 180 RU/ml will develop ESRD, as compared with participants with FGF-23 ≤ 180 RU/ml (about 9%). In the right panel of Figure 5, the gray lines, for the purpose of comparison, represent cumulative incidence



**Figure 5.** | Cumulative incidence functions for ESRD by FGF-23 groups on the basis of cause-specific hazards models with a time-dependent coefficient and subdistribution hazards models with a time-dependent coefficient. Note that the gray solid and dotted lines from subdistribution hazards models are cumulative incidence functions for ESRD by FGF-23 groups on the basis of subdistribution hazards models without time-dependent coefficient. FGF-23, fibroblast growth factor 23.

functions on the basis of the subdistribution hazards models without a time-dependent coefficient.

## Discussion

In this paper, we first introduce different types of events in survival analysis: an event of interest, competing events, and censoring events. Because of the dependence between the event of interest and competing events, standard analytical tools are not suitable. We then provide a review of two analytical tools, cause-specific hazards models and subdistribution hazards models, for analyzing time to the event of interest in the presence of competing events. Furthermore, a framework for determining an appropriate statistical approach in the setting of competing risks on the basis of the scientific goals of an analysis is provided. Finally, we use a real study of CKD to demonstrate the results and the interpretation of both cause-specific and subdistribution hazards models on the basis of association and prediction.

When hazard ratios from both cause-specific and subdistribution hazards are reported, under the proportionality assumptions, these hazard ratios are related to each other in specific ways (Supplemental Table 2). Therefore, we recommend readers to always examine the proportionality assumption and to correctly interpret the results for each model in published work.

Analytical strategies in the presence of competing events are not limited to cause-specific hazards models and subdistribution hazards models. Readers interested in an advanced treatment of this topic are encouraged to follow

up with the following references. Whereas the cause-specific hazards models are on the basis of the Cox proportional hazards model or technically semiparametric models, a more flexible modeling technique is proposed by Hinchliffe and Lambert (21) by relaxing the linear logarithm of time to the restricted cubic splines for logarithm of time in parametric Weibull models. One important limitation of subdistribution hazards models with time-dependent covariates is that the time-dependent covariate is not clearly defined after study subjects experience competing events. Many authors have pointed out this challenge in subdistribution hazards models and provided alternative approaches using a multistate models formulation (13,23–25). Another approach is to impute the occurrence of the event of interest for those who have experienced the competing events. For example, in a study of CKDs, if repeated measurements of eGFR are available, the time to ESRD for those who died could be estimated on the basis of progression of eGFR and a cutoff to define ESRD (26). Other authors have developed methods to address the dependence among the event of interest and competing events, such as mixed proportional hazards models in settings of a joint latent failure time distribution (10,27–29).

## Acknowledgments

Funding for the Chronic Renal Insufficiency Cohort Study was obtained under a cooperative agreement from the National Institute of Diabetes and Digestive and Kidney Diseases (grant nos. U01DK060990, U01DK060984, U01DK061022, U01DK061021,

U01DK061028, U01DK60980, U01DK060963, and U01DK060902). Additional funding was provided by CKD Biocon (U01DK85649).

#### Disclosures

None.

#### References

- Austin PC, Lee DS, Fine JP: Introduction to the analysis of survival data in the presence of competing risks. *Circulation* 133: 601–609, 2016
- Isakova T, Xie H, Yang W, Xie D, Anderson AH, Scialla J, Wahl P, Gutiérrez OM, Steigerwalt S, He J, Schwartz S, Lo J, Ojo A, Sondheimer J, Hsu CY, Lash J, Leonard M, Kusek JW, Feldman HI, Wolf M; Chronic Renal Insufficiency Cohort (CRIC) Study Group: Fibroblast growth factor 23 and risks of mortality and end-stage renal disease in patients with chronic kidney disease. *JAMA* 305: 2432–2439, 2011
- Bowe B, Xie Y, Xian H, Balasubramanian S, Al-Aly Z: Low levels of high-density lipoprotein cholesterol increase the risk of incident kidney disease and its progression. *Kidney Int* 89: 886–896, 2016
- Ricardo AC, Lash JP, Fischer MJ, Lora CM, Budoff M, Keane MG, Kusek JW, Martinez M, Nessel L, Stamos T, Ojo A, Rahman M, Soliman EZ, Yang W, Feldman HI, Go AS; CRIC and HCRIC Investigators: Cardiovascular disease among hispanics and non-hispanics in the chronic renal insufficiency cohort (CRIC) study. *Clin J Am Soc Nephrol* 6: 2121–2131, 2011
- Deo R, Shou H, Soliman EZ, Yang W, Arkin JM, Zhang X, Townsend RR, Go AS, Shlipak MG, Feldman HI: Electrocardiographic measures and prediction of cardiovascular and non-cardiovascular death in CKD. *J Am Soc Nephrol* 27: 559–569, 2016
- Feldman HI, Appel LJ, Chertow GM, Cifelli D, Cizman B, Daugirdas J, Fink JC, Franklin-Becker ED, Go AS, Hamm LL, He J, Hostetter T, Hsu CY, Jamerson K, Joffe M, Kusek JW, Landis JR, Lash JP, Miller ER, Mohler ER 3rd, Muntner P, Ojo AO, Rahman M, Townsend RR, Wright JT; Chronic Renal Insufficiency Cohort (CRIC) Study Investigators: The chronic renal insufficiency cohort (CRIC) study: Design and methods. *J Am Soc Nephrol* 14[Suppl 2]: S148–S153, 2003
- Lash JP, Go AS, Appel LJ, He J, Ojo A, Rahman M, Townsend RR, Xie D, Cifelli D, Cohan J, Fink JC, Fischer MJ, Gadegbeku C, Hamm LL, Kusek JW, Landis JR, Narva A, Robinson N, Teal V, Feldman HI; Chronic Renal Insufficiency Cohort (CRIC) Study Group: Chronic renal insufficiency cohort (CRIC) study: Baseline characteristics and associations with kidney function. *Clin J Am Soc Nephrol* 4: 1302–1311, 2009
- Fischer MJ, Go AS, Lora CM, Ackerson L, Cohan J, Kusek JW, Mercado A, Ojo A, Ricardo AC, Rosen LK, Tao K, Xie D, Feldman HI, Lash JP; CRIC and H-CRIC Study Groups: CKD in Hispanics: Baseline characteristics from the CRIC (Chronic renal insufficiency cohort) and Hispanic-CRIC studies. *Am J Kidney Dis* 58: 214–227, 2011
- Kaplan EL, Meier P: Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53: 457–481, 1958
- Cox DR: Regression models and life-tables. *J R Stat Soc B* 34: 187–220, 1972
- Prentice RL, Kalbfleisch JD, Peterson Jr. AV, Flournoy N, Farewell VT, Breslow NE: The analysis of failure times in the presence of competing risks. *Biometrics* 34: 541–554, 1978
- Lunn M, McNeil D: Applying Cox regression to competing risks. *Biometrics* 51: 524–532, 1995
- Putter H, Fiocco M, Geskus RB: Tutorial in biostatistics: Competing risks and multi-state models. *Stat Med* 26: 2389–2430, 2007
- Lim HJ, Zhang X, Dyck R, Osgood N: Methods of competing risks analysis of end-stage renal disease and mortality among people with diabetes. *BMC Med Res Methodol* 10: 97, 2010
- Fine JP, Gray RJ: A proportional hazards model for the sub-distribution of a competing risk. *J Am Stat Assoc* 94: 496–509, 1999
- Andersen PK, Geskus RB, de Witte T, Putter H: Competing risks in epidemiology: Possibilities and pitfalls. *Int J Epidemiol* 41: 861–870, 2012
- Lau B, Cole SR, Gange SJ: Competing risk regression models for epidemiologic data. *Am J Epidemiol* 170: 244–256, 2009
- Cheng SC, Fine JP, Wei LJ: Prediction of cumulative incidence function under the proportional hazards model. *Biometrics* 54: 219–228, 1998
- Beyersmann J, Dettenkofer M, Bertz H, Schumacher M: A competing risks analysis of bloodstream infection after stem-cell transplantation using subdistribution hazards and cause-specific hazards. *Stat Med* 26: 5360–5369, 2007
- Muñoz A, Abraham AG, Matheson M, Wada N: Non-proportionality of hazards in the competing risks framework. In: *Risk Assessment and Evaluation of Predictions*, New York, Springer, 2013, pp 3–22
- So Y, Lin G, Johnson G: Using the PHREG procedure to analyze competing-risks data. Presented at the SAS Global Forum, Washington, DC, March 23–26, 2014. Available at: <http://support.sas.com/rnd/app/stat/papers/2014/competingrisk2014.pdf>
- Hinchliffe SR, Lambert PC: Flexible parametric modelling of cause-specific hazards to estimate cumulative incidence functions. *BMC Med Res Methodol* 13: 13, 2013
- Latouche A, Porcher R, Chevret S: A note on including time-dependent covariate in regression model for competing risks data. *Biometrical Journal* 47: 807–814, 2005
- Latouche A, Boisson V, Chevret S, Porcher R: Misspecified regression model for the subdistribution hazard of a competing risk. *Statistics in Medicine* 26: 965–974, 2007
- Beyersmann J, Schumacher M: Time-dependent covariates in the proportional subdistribution hazards model for competing risks. *Biostatistics* 9: 765–776, 2008
- Lu JL, Kalantar-Zadeh K, M JZ, Quarles LD, Kovesdy CP: Association of body mass index with outcomes in patients with CKD. *J Am Soc Nephrol* 25: 2088–2096, 2014
- Cox DR: Regression models and life-tables. *J R Stat Soc B* 34: 187–220, 1972
- Tsiatis A: A nonidentifiability aspect of the problem of competing risks. *Proc Natl Acad Sci USA* 72: 20–22, 1975
- Abbring JH, Van Den Berg GJ: The identifiability of the mixed proportional hazards competing risks model. *J R Stat Soc Series B Stat Methodol* 65: 701–710, 2003

Published online ahead of print. Publication date available at [www.cjasn.org](http://www.cjasn.org).

This article contains supplemental material online at <http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.10301016/-/DCSupplemental>.