KIDNEY RESEARCH AND CLINICAL PRACTICE

Effect of estimating equations for glomerular filtration rate on novel surrogate markers for renal outcome

Kipyo Kim[1], Eunji Baek[2], Suryeong Go[2], Hyung-Eun Son[2], Ji-Young Ryu[2], Yongjin Yi[3], Jong Cheol Jeong[2], Sejoong Kim[2,4], Ho Jun Chin[2,4]

[1]Division of Nephrology and Hypertension, Department of Internal Medicine, Inha University College of Medicine, Incheon, Republic of Korea
[2]Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea
[3]Division of Nephrology, Department of Internal Medicine, Dankook University Hospital, Dankook University College of Medicine, Cheonan, Republic of Korea
[4]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

Backgrounds: Recently, alternative surrogate endpoints such as a 30% or 40% decline in estimated glomerular filtration rate (eGFR) or eGFR slope over 2 to 3 years have been proposed for predicting renal outcomes. However, the impact of GFR estimation methods on the accuracy and effectiveness of surrogate markers is unknown.

Methods: We retrospectively enrolled participants in health screening programs at three hospitals from 1995 to 2009. We defined two different participant groups as YR1 and YR3, which had available 1-year or 3-year eGFR values along with their baseline eGFR levels. We compared the effectiveness of eGFR percentage change or slope to estimate end-stage renal disease (ESRD) risk according to two estimating equations (modified Modification of Diet in Renal Disease equation [eGFRm] and Chronic Kidney Disease-Epidemiology Collaboration (CKD-EPI) equation [eGFRc]) for GFR.

Results: In the YR1 and YR3 groups, 9,971 and 10,171 candidates were enrolled and ESRD incidence during follow-up was 0.26% and 0.19%, respectively. The eGFR percentage change was more effective than eGFR slope in estimating ESRD risk, regardless of the method of estimation. A 40% of decline in eGFR was better than 30%, and a 3-year baseline period was better than a 1-year period for prediction accuracy. Although some diagnostic indices from the CKD-EPI equation were better, we found no significant differences in the discriminative ability and hazard ratios for incident ESRD between eGFRc and eGFRm in either eGFR percentage change or eGFR slope.

Conclusion: There were no significant differences in the prediction accuracy of GFR percentage change or eGFR slope between eGFRc and eGFRm in the general population.

Keywords: Chronic kidney disease, End-stage renal disease, Glomerular filtration rate, Renal endpoint, Surrogate endpoint

Introduction

Chronic kidney disease (CKD) has been recognized as a leading cause of morbidity and mortality [1,2]. Nevertheless, current therapies for preventing CKD progression are still lacking [3], which is largely attributable to the absence of well-validated surrogate endpoints for renal outcomes [4]. Established endpoints for CKD progression, such as doubling of serum creatinine or end-stage renal disease (ESRD), are late events indicating substantial kidney damage. Furthermore, such studies require long-term follow-up and a large number of participants to produce reliable data. Considering these challenges, several alternative surrogate markers for CKD progression have been proposed [5]. In the 2012 scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration, a decline of 30% or 40% in the estimated glomerular filtration rate (eGFR) over 2 to 3 years was proposed as an acceptable alternative surrogate endpoint for clinical trials involving CKD [6]. While these newly-suggested endpoints have been utilized in some studies [7,8], experts recommend their careful application in clinical trials [6]. In particular, surrogate endpoints based on percentage changes in eGFR have been mainly validated in the CKD population and are less applicable in patients with a high baseline GFR [9]. Surrogate markers could prove more useful if they are validated in patients with early-stage CKD who have high GFRs. Therefore, different candidate surrogate endpoints such as GFR slope or changes in albuminuria have been evaluated in recent studies [9–12]. However, the efficacy of these endpoints needs to be further validated. Moreover, differences in the efficacy of novel surrogate endpoints based on the GFR-estimating equations are not well-studied.

Currently, the Modification of Diet in Renal Disease (MDRD) and Chronic Kidney Disease-Epidemiology Collaboration (CKD-EPI) equations are widely used to estimate GFR [13,14]. The later-developed CKD-EPI equation is increasingly used due to better accuracy at GFRs of >60 mL/min/1.73 m^2. However, none of these equations are optimally applicable across all GFR ranges and in different populations [15]. The eGFR decline was noninferior to the measured GFR decline in predicting renal outcomes [16]; however, the specific GFR estimation method which may improve predictive accuracy when utilizing novel surrogate endpoints is unknown. Application of different equations

could lead to misclassification of CKD [17] or to variability in the calculation of GFR decline. Therefore, we aimed to evaluate the impact of GFR estimation methods on the efficacy of GFR decline or GFR slope for predicting renal outcomes in the general population.

Methods

Study design

We retrospectively enrolled 143,890 participants aged ≥18 years with eGFR of ≥15 mL/min/1.73 m^2 who underwent routine health screenings between May 1995 and April 2009 at one of three university-affiliated hospitals (Seoul National University Hospital, Seoul National University Bundang Hospital, and Seoul National University Boramae Medical Center) in Korea. Clinical characteristics including age, sex, medical history, and laboratory values, were collected from the electronic medical records of each participant at each hospital. The eGFR was calculated using the modified MDRD equation (eGFRm) or the CKD-EPI 2009 equation (eGFRc) using isotope dilution mass spectrometry-traceable creatinine values [13,14]. To determine eGFR decline and slope, we defined two different baseline period groups, YR1 and YR3, which consisted of participants with available 1-year or 3-year eGFR values along with their baseline eGFR levels. The development of ESRD by December 2017 was determined using the ESRD registry of the Korean Society of Nephrology. The study was approved by the Institutional Review Board (IRB) of Seoul National University Hospital (No. B-1801/442-003). The requirement for written informed consent was waived by the IRB, and all data were fully anonymized before analysis.

Statistical analyses

Data were presented as percentages for categorical variables and as mean ± standard deviation for continuous variables. Each variable was compared using the t-test for continuous variables and the chi-square or Fisher exact test for categorical variables. Receiver operating characteristic (ROC) curve analyses were performed to examine the discriminant ability of each surrogate marker for predicting ESRD development. The Bland-Altman plot was also used to represent the agreement of eGFR percentage change and eGFR slope according

to the equations used for GFR estimation. The areas under the ROC curves (AUCs) of different surrogate markers were compared using the DeLong test.

The performance of surrogate markers was also presented in a manner similar to that of diagnostic tests [18]. To evaluate overall performance, diagnostic accuracy was shown by using paired indicators such as sensitivity and specificity, positive predictive value (PPV) and negative predictive value, and positive likelihood ratio and negative likelihood ratio. In particular, likelihood ratios are useful measures from a clinical perspective and are directly linked to posttest probabilities of event occurrence [19]. Overall diagnostic accuracy and diagnostic odds ratios (DOR) were used as global measures. DORs are not dependent on disease prevalence [20] and are therefore more useful for comparing accuracy between surrogate markers for rare events. Comparisons of sensitivity and specificity achieved with different threshold values were performed using the McNemar test, and positive and negative predictive values were compared using a weighted generalized score statistic [21].

Multivariable Cox proportional hazards analysis with a restricted cubic spline of three knots was performed to examine the association of GFR decline and GFR slope with renal survival. Hazard ratios (HRs) were adjusted by age, sex, and factors related to incident ESRD including systolic blood pressure, diabetes mellitus, eGFR, uric acid, albumin, alkaline phosphatase, glucose, hemoglobin, and urine dipstick protein results. The completeness of these variables is shown in Supplementary Table 1 (available online), and listwise deletion was used in the multivariable analyses. The p-values of <0.05 were considered statistically significant. All analyses were performed using IBM SPSS version 23 (IBM Corp., Armonk, NY, USA) and R software version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Baseline characteristics

Of 143,890 subjects aged ≥18 years with eGFRc of ≥15 mL/min/1.73 m^2, we finally included 9,972 and 10,171 individuals in the YR1 and YR3 groups, respectively, after excluding patients with either missing eGFRc values or inadequate follow-up data (Fig. 1). The YR1 and YR3 groups were followed for 180.1 ± 38.6 months and 178.8 ± 33.4 months, respective-

ly. The mean age of the YR1 group was 54 years, 60.0% were male, the mean eGFRc was 95.8 mL/min/1.73 m^2, and 10.6% were diabetic (Table 1). The mean age of the YR3 group was 53 years, 59.0% were male, the mean eGFRc was 95.9 mL/min/1.73 m^2, and 10.4% were diabetic. In total, 1.8% and 2.3% of YR1 participants and 1.7% and 2.3% of YR3 patients had an eGFRc of <60 mL/min/1.73 m^2 and an eGFRm of <60 mL/min/1.73 m^2, respectively. The YR1 and YR3 groups did not differ with respect to sex, presence of diabetes, blood pressure, or laboratory values.

At baseline, eGFRc showed a greater mean value compared to eGFRm in both the YR1 (95.8 mL/min/1.73 m^2 vs. 95.4 mL/min/1.73 m^2) and YR3 (95.9 mL/min/1.73 m^2 vs. 95.0 mL/min/1.73 m^2) groups. The standard deviation and intersubject coefficient of variation (CV) were greater for eGFRm, as these values were more wildly distributed across the range of eGFR than eGFRc values were (Supplementary Table 2, available online). The Bland-Altman plot showed how differences between eGFRc and eGFR in both the YR1 and YR3 groups were prominent at high GFR levels (≥90 mL/min/1.73 m^2) (Supplementary Fig. 1, available online).
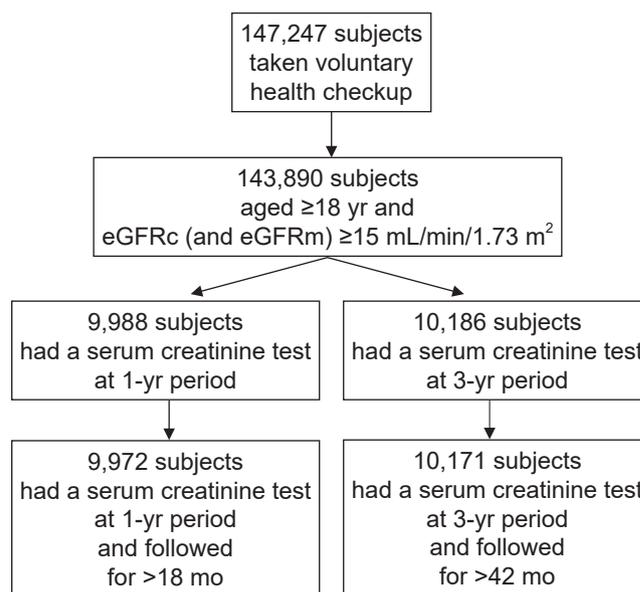
Figure 1. Selection of study participants. 1-year period, 6–18 months after the first examination; 3-year period, 30–42 months after the first examination.
eGFR, estimated glomerular filtration rate; eGFRc, eGFR by the 2009 Chronic Kidney Disease-Epidemiology Collaboration creatinine equation; eGFRm, eGFR by the modified Modification of Diet in Renal Disease equation.

Table 1. Baseline characteristics of the study participants

Variable	Participant		p-value[a]
	YR1	YR3	
No. of patients	9,972	10,171	
Age (yr)	53.8 ± 11.1	53.3 ± 11.2	0.004
Male sex	5,980 (60.0)	5,998 (59.0)	0.15
Diabetes mellitus	1,053 (10.6)	1,056 (10.4)	0.98
Hypertension	2,539 (25.5)	2,513 (24.7)	0.25
Body mass index (kg/m^2)	24.0 ± 2.9	23.9 ± 2.9	0.09
SBP (mmHg)	121 ± 17	120 ± 17	0.14
DBP (mmHg)	75 ± 12	75 ± 12	0.83
Hemoglobin (g/dL)	14.6 ± 1.5	14.5 ± 1.5	0.28
Cholesterol (mg/dL)	205 ± 36	204 ± 36	0.32
Triglyceride (mg/dL)	129 ± 81	128 ± 87	0.43
HDL-C (mg/dL)	56 ± 14	56 ± 14	0.18
Fasting glucose (mg/dL)	99 ± 24	99 ± 24	0.48
Protein (g/dL)	7.4 ± 0.4	7.4 ± 0.4	0.27
Albumin (g/dL)	4.4 ± 0.3	4.4 ± 0.3	0.38
AST (U/L)	26 ± 15	26 ± 18	0.59
ALT (U/L)	29 ± 28	29 ± 27	0.86
ALP (U/L)	69 ± 21	69 ± 20	0.50
Uric acid (mg/dL)	5.5 ± 1.4	5.5 ± 1.5	0.69
Calcium (mg/dL)	9.1 ± 0.5	9.1 ± 0.5	0.21
Phosphorus (mg/dL)	3.7 ± 0.8	3.7 ± 0.8	0.98
eGFRc (mL/min/1.73 m^2)	95.8 ± 15.4	95.9 ± 15.4	0.39
≥90	6,773 (67.9)	6,949 (68.3)	
<90, ≥60	3,024 (30.3)	3,049 (30.0)	
<60, ≥30	162 (1.6)	161 (1.6)	
<30	13 (0.1)	12 (0.1)	
eGFRm (mL/min/1.73 m^2)	95.4 ± 27.3	95.0 ± 25.9	0.31
≥90	5,148 (51.6)	5,233 (51.5)	
<90, ≥60	4,592 (46.0)	4,702 (46.2)	
<60, ≥30	219 (2.2)	224 (2.2)	
<30	13 (0.1)	12 (0.1)	
Proteinuria by dipstick			0.29
None or trace	8,546 (85.7)	8,832 (87.2)	
1+	1,102 (11.1)	1,010 (10.0)	
≥2+	288 (2.9)	283 (2.8)	

Data are expressed as number only, mean ± standard deviation, or number (%).

ALP, alkaline phosphatase; ALT, alanine transaminase; AST, aspartate transaminase; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; eGFRc, eGFR by the 2009 Chronic Kidney Disease-Epidemiology Collaboration creatinine equation; eGFRm, eGFR by the modified Modification of Diet in Renal Disease equation; HDL-C, high density lipoprotein cholesterol; SBP, systolic blood pressure; YR1, 1-year eGFR group; YR3, 3-year eGFR group.

[a]Comparison between YR1 and YR3.

During the baseline period, percentage changes in eGFR differed between eGFRc and eGFRm in YR1, but not in YR3 (Supplementary Table 2). Likewise, while the eGFRc slope was steeper than the eGFRm slope in YR1, this difference in inclination was not observed in YR3. The CV for percentage change in eGFR and for the eGFR slope was much greater for eGFRm compared to eGFRc in both the YR1 and YR3 groups. The Bland-Altman plot showed that the larger the absolute value of eGFR percentage change and eGFR slope, the larger the difference between eGFRc and eGFRm. In particular, subjects with high eGFRs showed greater differences in eGFR percentage change and eGFR slope between the CKD-EPI and MDRD equations.

Prediction accuracy for ESRD according to surrogate markers

During the follow-up period, the rate of ESRD incidence was 0.26% and 0.19% in the YR1 and YR3 groups, respectively. Surrogate markers showed better discrimination for ESRD development in the YR3 group compared to the YR1 group (Table 2). ROC analysis showed that ESRD prediction accuracy was significantly greater when using percentage change in eGFR compared to eGFR slope (Fig. 2, Table 2). Among the studied surrogate markers, the percentage change in eGFRc in YR3 showed the highest AUC (0.837), while eGFRc slope in YR1 showed the lowest AUC (0.632). AUCs for percentage changes in eGFRc and eGFRm, and those for yearly slopes for eGFRc and eGFRm values, did not differ in either the YR1 or YR3 groups (Table 2). When diagnostic indicators were evaluated based on percentage changes in eGFR and the estimating equations, changes of ≥30% in eGFRc and eGFRm showed an acceptable specificity of >90% and a high DOR (Table 3). Although the PPV of each endpoint was relatively small due to the low incidence of ESRD, application of a higher threshold yielded a greater PPV. Overall, the diagnostic indices according to the two equations revealed similar patterns. However, the eGFRc criteria had higher DOR, specificity, and PPV than the eGFRm criteria at all eGFR thresholds in both the YR1 and YR3 groups (Table 3). In particular, the PPV and specificity of a 30% or 40% decline in eGFR were statistically different based on the GFR- estimating equation used (Supplementary Table 3, available online). The PPV associated with a 40% decline in eGFR was not statistically different compared to conventional endpoints (57% eGFRc or 55% eGFRm decline) in YR3, but not for a 30% decline or less. Subgroup analyses showed that 30% and 40% declines in eGFR showed higher DORs in patients with age

Table 2. Comparison of AUCs of ROC curves for eGFR changes to estimate incident ESRD

Period	Variable	Percent of eGFRm change	Slope of eGFRc per year	Slope of eGFRm per year
YR1	Percent of eGFRc change (AUC, 0.71)	0.20	0.004	0.05
	Percent of eGFRm change (AUC, 0.68)		0.001	0.03
	Slope of eGFRc per year (AUC, 0.63)			0.72
YR3	Percent of eGFRc change (AUC, 0.84)	0.08	<0.001	0.001
	Percent of eGFRm change (AUC, 0.80)		0.004	0.002
	Slope of eGFRc per year (AUC, 0.66)			0.13

AUC, area under the ROC curve; eGFR, estimated glomerular filtration rate; eGFRc, eGFR by the 2009 Chronic Kidney Disease-Epidemiology Collaboration creatinine equation; eGFRm, eGFR by the modified Modification of Diet in Renal Disease equation; ESRD, end-stage renal disease; ROC, receiver operating characteristic; YR1, 1-year eGFR group; YR3, 3-year eGFR group.
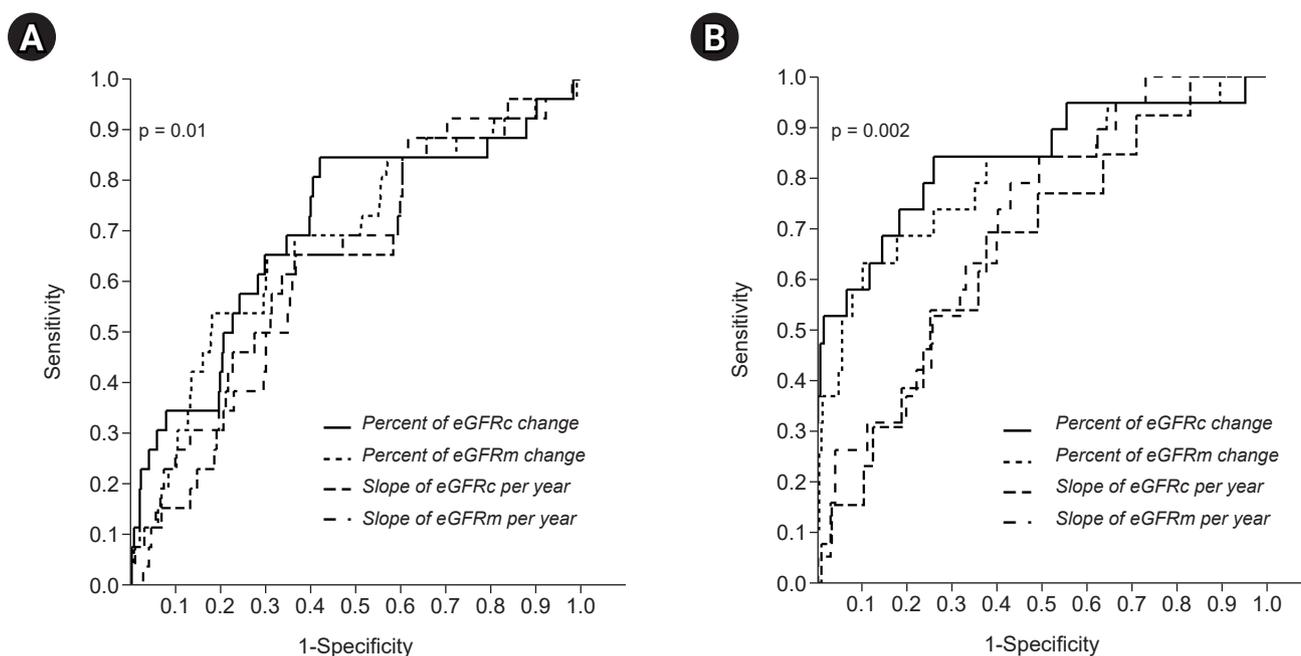
Figure 2. Receiver operating characteristic curves for changes in eGFR. During 1-year (A) and 3-year (B) periods for ESRD estimation. The eGFR was estimated based on isotope dilution mass spectrometry-traceable creatinine.
eGFR, estimated glomerular filtration rate; eGFRc, eGFR by the 2009 Chronic Kidney Disease-Epidemiology Collaboration creatinine equation; eGFRm, eGFR by the modified Modification of Diet in Renal Disease equation; ESRD, end-stage renal disease.

of <65 years and proteinuria grade >1+. The eGFRc values demonstrated a better DOR than eGFRm values in patients aged <65 years and in those with eGFR of ≥60 mL/min/1.73 m^2 (Supplementary Table 4, 5; available online).

Associations between changes in eGFR and ESRD risk

In the restricted cubic spline model, percentage decline in eGFR was not significantly associated with the risk for ESRD in YR1 (Fig. 3). However, in YR3, a greater percentage decline in eGFR was significantly associated with a higher

risk for developing ESRD at all ranges. These associations between changes in eGFR and HR were comparable using eGFRc and eGFRm. A negative eGFR slope was associated with a modest increase in the adjusted HR for ESRD in YR1 (Fig. 4), while this correlation with eGFR slope was stronger in YR3. While eGFRm slope had a higher estimated HR than eGFRc, the former demonstrated a wider confidence interval.

In subgroup analyses stratified by age, sex, presence of diabetes, eGFR, and presence of proteinuria, adjusted HRs for percentage change in eGFR for each criterion were not

Table 3. Diagnostic accuracy of the criteria for eGFR changes to estimate ESRD

Variable	Criteria for eGFRc change (No. of subjects at risk)			Criteria for eGFRm change (No. of subjects at risk)		
YR1 group (n = 9,972)	30.0% (n = 139)	40.0% (n = 39)	57.0% (n = 5)	30.0% (n = 661)	40.0% (n = 261)	55.0% (n = 32)
Sensitivity	0.115 (0.024–0.302)	0.077 (0.009–0.251)	0.038 (0.001–0.196)	0.154 (0.044–0.349)	0.115 (0.024–0.302)	0.038 (0.001–0.196)
Specificity	0.986 (0.984–0.989)	0.996 (0.995–0.997)	1.000 (0.999–1.000)	0.934 (0.929–0.939)	0.974 (0.971–0.977)	0.997 (0.996–0.998)
PPV	0.022 (0.004–0.062)	0.051 (0.006–0.173)	0.200 (0.005–0.716)	0.006 (0.002–0.015)	0.011 (0.002–0.033)	0.031 (0.001–0.162)
NPV	0.998 (0.996–0.999)	0.998 (0.996–0.998)	0.997 (0.996–0.998)	0.998 (0.996–0.999)	0.998 (0.996–0.998)	0.997 (0.996–0.998)
LR (+)	8.438 (2.873–24.782)	20.678 (5.255–81.360)	95.635 (11.060–826.976)	2.329 (0.943–5.754)	4.448 (1.524–12.982)	12.340 (1.749–87.062)
LR (–)	0.897 (0.781–1.030)	0.927 (0.829–1.035)	0.962 (0.891–1.039)	0.906 (0.769–1.067)	0.908 (0.790–1.043)	0.965 (0.893–1.042)
DOR	9.409 (2.792–31.709)	22.318 (5.090–97.861)	99.420 (10.731–921.066)	2.571 (0.883–7.482)	4.899 (1.461–16.416)	12.794 (1.681–97.376)
Accuracy	0.984 (0.981–0.986)	0.994 (0.992–0.995)	0.997 (0.996–0.998)	0.932 (0.927–0.937)	0.972 (0.968–0.975)	0.994 (0.927–0.996)
YR3 group (n = 10,171)	30.0% (n = 186)	40.0% (n = 49)	57.0% (n = 6)	30.0% (n = 820)	40.0% (n = 245)	55.0% (n = 26)
Sensitivity	0.526 (0.289–0.756)	0.368 (0.163–0.616)	0.105 (0.013–0.331)	0.526 (0.289–0.756)	0.368 (0.163–0.616)	0.105 (0.013–0.331)
Specificity	0.983 (0.980–0.985)	0.996 (0.994–0.997)	1.000 (0.999–1.000)	0.920 (0.915–0.925)	0.977 (0.973–0.979)	0.998 (0.996–0.998)
PPV	0.054 (0.026–0.097)	0.143 (0.059–0.272)	0.333 (0.043–0.777)	0.012 (0.006–0.022)	0.029 (0.012–0.058)	0.077 (0.009–0.251)
NPV	0.999 (0.998–1.000)	0.999 (0.998–0.999)	0.998 (0.997–0.999)	0.999 (0.998–1.000)	0.999 (0.998–0.999)	0.998 (0.997–0.999)
LR (+)	30.359 (19.338–47.660)	89.053 (45.954–172.570)	267.158 (51.999–1372.600)	6.596 (4.284–10.157)	15.715 (8.608–28.691)	44.526 (11.309–175.312)
LR (–)	0.482 (0.300–0.774)	0.634 (0.450–0.894)	0.895 (0.767–1.044)	0.515 (0.320–0.827)	0.647 (0.459–0.912)	0.897 (0.769–1.046)
DOR	62.980 (25.280–156.899)	140.417 (52.691–374.200)	298.471 (51.207–1739.690)	12.815 (5.192–31.627)	24.299 (9.483–62.266)	49.647 (10.870–226.754)
Accuracy	0.982 (0.979–0.984)	0.995 (0.993–0.996)	0.998 (0.997–0.999)	0.919 (0.914–0.925)	0.975 (0.972–0.978)	0.996 (0.995–0.997)

Data are expressed as eGFR change (%) or index (95% confidence interval).

DOR, diagnostic odds ratio; eGFR, estimated glomerular filtration rate; eGFRc, eGFR by the 2009 Chronic Kidney Disease-Epidemiology Collaboration creatinine equation; eGFRm, eGFR by the modified Modification of Diet in Renal Disease equation; ESRD, end-stage renal disease; LR (–), negative likelihood ratio; LR (+), positive likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; YR1, 1-year eGFR group; YR3, 3-year eGFR group.
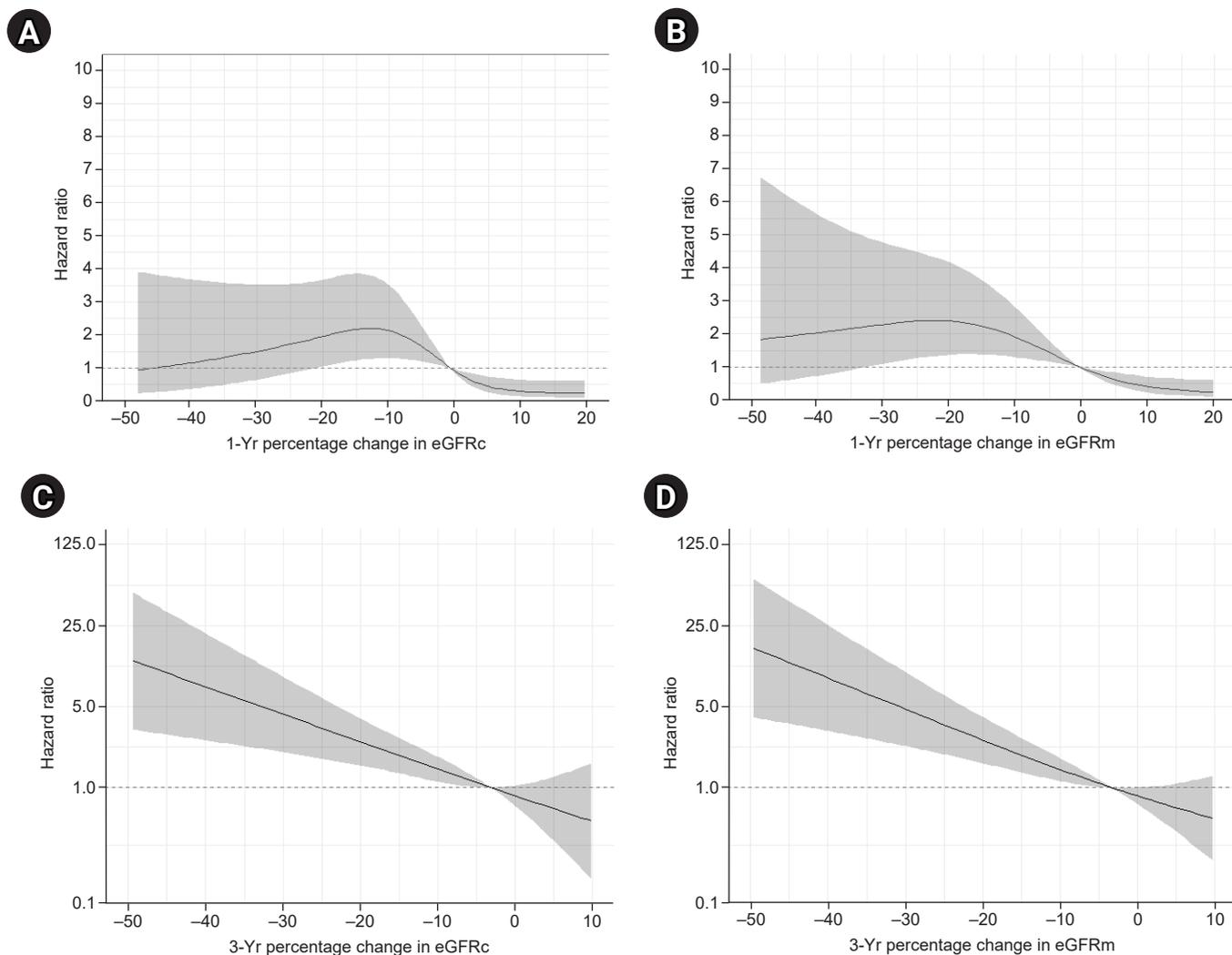
Figure 3. Adjusted hazard ratio for end-stage renal disease associated with percent change in eGFR. eGFRc (A) and eGFRm (B) over 1 year; eGFRc (C) and eGFRm (D) over 3 years.
eGFR, estimated glomerular filtration rate; eGFRc, eGFR by the 2009 Chronic Kidney Disease-Epidemiology Collaboration creatinine equation; eGFRm, eGFR by the modified Modification of Diet in Renal Disease equation.

significant in YR1, except for eGFR of ≥60 mL/min/1.73 m^2 (Supplementary Table 6, available online). Meanwhile, percentage changes in eGFR in YR3 were associated with significantly higher HRs for each subgroup, especially in young patients and those with proteinuria. Overall, estimated HRs based on eGFRc or eGFRm were similar, though HRs estimated using eGFRc were 2 to 3 times greater than those derived using eGFRm in patients with eGFR of ≥60 mL/min/1.73 m^2.

Discussion

In this study, we evaluated the predictive accuracy of several novel surrogate markers for ESRD, including eGFR percentage change and eGFR slope over 1- and 3-year baseline periods in the general population, using different estimating equations. Our findings showed that a 30% or 40% decline in eGFR may be a more accurate surrogate endpoint than the eGFR slope when predicting renal outcome. We also showed that these values are more predictive over a baseline period of 3 years compared to 1 year. Percentage changes in eGFRc
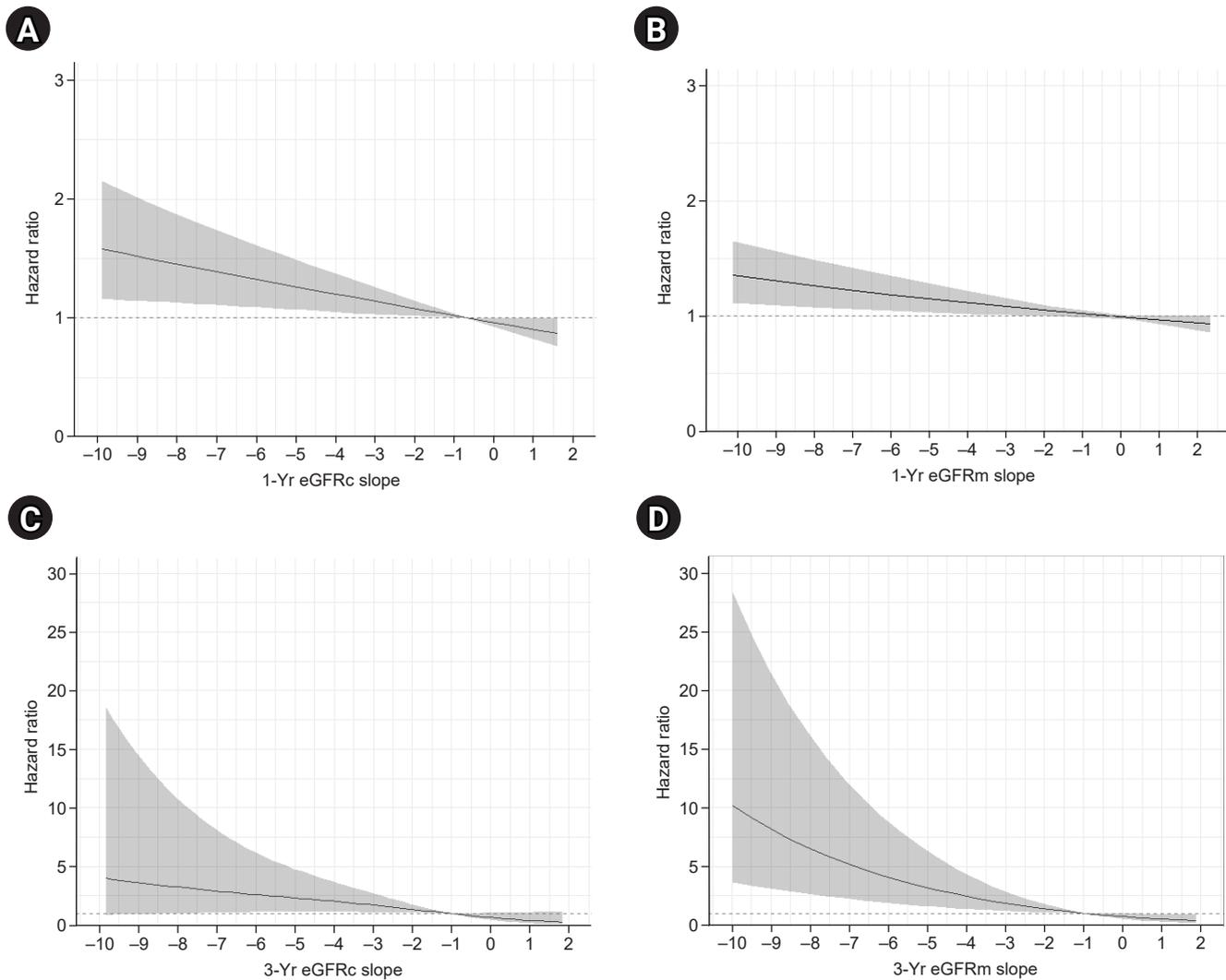
Figure 4. Adjusted hazard ratio for end-stage renal disease associated with eGFR slope. eGFRc slope (A) and eGFRm slope (B) over 1 year; the eGFRc slope (C) and eGFRm slope (D) over 3 years.
eGFR, estimated glomerular filtration rate; eGFRc, eGFR by the 2009 Chronic Kidney Disease-Epidemiology Collaboration creatinine equation; eGFRm, eGFR by the modified Modification of Diet in Renal Disease equation.

showed higher specificity and PPV compared to those in eG-FRm and were more likely to be associated with the development of ESRD in patients with eGFR of ≥60 mL/min/1.73 m². Nevertheless, there were no significant differences in discriminative ability (AUC) or adjusted HRs between eG-FRc and eGFRm for predicting ESRD risk.

Recent studies have provided evidence in favor of using alternative surrogate endpoints such as eGFR percentage decline and eGFR slope for predicting renal outcomes [10–12,22,23]. Indeed, an ongoing randomized trial to evaluate the effect of sodium-glucose cotransporter-2 inhibitors on renal outcomes set eGFR decline of ≥40% as one of the primary endpoints [24]. Nevertheless, novel surrogate endpoints have not been well validated in the general population. Overall, our findings are similar to those of previous studies in the CKD population showing that eGFR changes over 1 year were insufficient to predict renal outcomes. Researchers have recommended the use of surrogate markers over 2 to 3 years [6]. Furthermore, a 40% decline in eGFR is more acceptable than a 30% decline [25], which was also consistent with our results. We found that a 40% decline in eGFR over 3 years was the most reliable candidate surrogate endpoint in this study. Since our study was based on the Korean population, our findings highlight the significance

of these early surrogate markers in Asian populations. While the eGFR slope showed moderate predictive accuracy as well, it was lower than that achieved with percentage change in eGFR and was a comparatively weaker surrogate marker. These findings may be related to characteristics of eGFR slope such as greater variability in eGFR at higher levels, nonlinear trajectories, and nonuniform distributions [26].

These early-identifiable markers are important, as conventional endpoints such as creatinine doubling or ESRD require much longer follow-up periods in the general population. The application of these surrogate endpoints can enable early detection and management of CKD progression. However, there are some considerations regarding surrogate markers of eGFR change in patients with high baseline GFR levels. The MDRD and CKD-EPI equations are most commonly used to assess kidney function in medical research and clinical practice. Clinically meaningful differences between the two equations are mainly seen when GFR levels are high [27]. The CKD-EPI equation shows greater accuracy than the MDRD equation when estimating GFR in individuals with normal kidney function [14]. Thus, using the CKD-EPI equation can decrease the chances of misclassifying patients at low risk of ESRD [28]. Since misclassification of at-risk patients affects the efficacy of surrogate endpoints, it is clinically important to assess their accuracy based on the equation used for GFR estimation. Our findings demonstrated that both the CKD-EPI and MDRD equations were acceptable at estimating the risk for ESRD in the general population based on the results of the ROC and Cox proportional hazards analyses. These results support the application of both equations for predicting renal outcome and alleviate concerns about the use of the MDRD equation in utilizing novel surrogate endpoints.

However, we found that the CKD-EPI equation was better with respect to some diagnostic indices such as specificity, DOR, and PPV, representing more accurate prediction for true endpoint. Our findings can be attributed to (1) less bias in estimating GFR and (2) lower within-subject variability of CKD-EPI equation in patients with high GFR [14,29]. As shown in our Bland-Altman plot analysis (Supplementary Fig. 1), the differences between eGFRc and eGFRm increased at higher GFRs, and greater differences in both eGFR percentage change and slope were observed mainly in patients with higher GFRs. These findings suggest that the greater bias associated with the MDRD equation could

affect its predictive accuracy. In addition, our results may be indicative of the high within-subject variability of the MDRD equation. Within-subject variability was usually higher at higher GFRs and was significantly higher in the application of the MDRD equation compared to the CKD-EPI equation [29]. Changes in eGFR between two time points may be partly affected by this within-subject variability.

There are certain limitations in this study. The number of enrolled patients was relatively small, and the rate of incidence of ESRD was low due to the characteristics of the studied population. Surrogate markers based on changes in eGFR were assessed using serum creatinine measurements merely at the beginning and the end of the defined baseline period. Therefore, GFR variability and acute treatment effects during the baseline period were not considered. In addition, several unmeasured confounders could be present, considering the observational nature of the study. Nevertheless, our study rigorously analyzed the differences in different novel surrogate endpoints with respect to the equations used for estimating eGFR, which allowed us to provide detailed implications. To the best of our knowledge, the impact of GFR-estimating equations on the efficacy of novel surrogate markers for predicting renal outcomes has not been evaluated previously.

In conclusion, there were no significant differences in estimated ESRD risk using the GFR percentage change or eGFR slope between the CKD-EPI equation and the modified MDRD equation in the general population. Surrogate markers using the CKD-EPI equation may be slightly more accurate in patients with high GFRs.

Conflict of interest

All authors have no conflicts of interest to declare.

Authors' contributions

Conceptualization: HJC
Data curation: KK, EB, SG, HES, JYR, YY, JCJ, SK
Formal analysis: KK, YY, HJC
Investigation: EB, SG, HES, JYR
Writing–original draft: KK
Writing–review & editing: All authors
All authors read and approved the final manuscript.

ORCID

Kipyo Kim, https://orcid.org/0000-0003-4166-1801
Eunji Baek, https://orcid.org/0000-0001-9226-7703
Suryeong Go, https://orcid.org/0000-0001-8118-7994
Hyung-Eun Son, https://orcid.org/0000-0002-8719-3823
Ji-Young Ryu, https://orcid.org/0000-0003-4134-1007
Yongjin Yi, https://orcid.org/0000-0001-8553-7189
Jong Cheol Jeong, https://orcid.org/0000-0003-0301-7644
Sejoong Kim, https://orcid.org/0000-0002-7238-9962
Ho Jun Chin, https://orcid.org/0000-0002-3710-0190

References

1. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N Engl J Med* 2004;351:1296–1305.

2. Chronic Kidney Disease Prognosis Consortium; Matsushita K, van der Velde M, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *Lancet* 2010;375:2073–2081.

3. Palmer SC, Sciancalepore M, Strippoli GF. Trial quality in nephrology: how are we measuring up? *Am J Kidney Dis* 2011;58: 335–337.

4. Ferguson TW, Komenda P, Tangri N. Change in estimated glomerular filtration rate and outcomes in chronic kidney disease. *Curr Opin Nephrol Hypertens* 2016;25:240–244.

5. Weldegiorgis M, de Zeeuw D, Heerspink HJ. Renal end points in clinical trials of kidney disease. *Curr Opin Nephrol Hypertens* 2015;24:284–289.

6. Levey AS, Inker LA, Matsushita K, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. *Am J Kidney Dis* 2014;64:821–835.

7. Perkovic V, Koitka-Weber A, Cooper ME, et al. Choice of endpoint in kidney outcome trials: considerations from the EMPA-REG OUTCOME® trial. *Nephrol Dial Transplant* 2020;35:2103–2111.

8. Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes. *N Engl J Med* 2017;377:644–657.

9. Levey AS, Gansevoort RT, Coresh J, et al. Change in albuminuria and GFR as end points for clinical trials in early stages of CKD: a scientific workshop sponsored by the National Kidney Foundation in Collaboration with the US Food and Drug Ad-

ministration and European Medicines Agency. *Am J Kidney Dis* 2020;75:84–104.

10. Grams ME, Sang Y, Ballew SH, et al. Evaluating glomerular filtration rate slope as a surrogate end point for ESKD in clinical trials: an individual participant meta-analysis of observational data. *J Am Soc Nephrol* 2019;30:1746–1755.

11. Greene T, Ying J, Vonesh EF, et al. Performance of GFR slope as a surrogate end point for kidney disease progression in clinical trials: a statistical simulation. *J Am Soc Nephrol* 2019;30:1756–1769.

12. Inker LA, Heerspink HJ, Tighiouart H, et al. GFR slope as a surrogate end point for kidney disease progression in clinical trials: a meta-analysis of treatment effects of randomized controlled trials. *J Am Soc Nephrol* 2019;30:1735–1745.

13. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med* 1999;130:461–470.

14. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604–612.

15. Earley A, Miskulin D, Lamb EJ, Levey AS, Uhlig K. Estimating equations for glomerular filtration rate in the era of creatinine standardization: a systematic review. *Ann Intern Med* 2012;156:785–795.

16. Ku E, Xie D, Shlipak M, et al. Change in measured GFR versus eGFR and CKD outcomes. *J Am Soc Nephrol* 2016;27:2196–2204.

17. Rocco MV, Chapman A, Chertow GM, et al. Chronic kidney disease classification in systolic blood pressure intervention trial: comparison using modification of diet in renal disease and CKD-Epidemiology Collaboration Definitions. *Am J Nephrol* 2016;44:130–140.

18. Rector TS, Taylor BC, Wilt TJ. Chapter 12: systematic review of prognostic tests. *J Gen Intern Med* 2012;27:S94–S101.

19. Eusebi P. Diagnostic accuracy measures. *Cerebrovasc Dis* 2013;36:267–272.

20. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 2003;56:1129–1135.

21. Kosinski AS. A weighted generalized score statistic for comparison of predictive values of diagnostic tests. *Stat Med* 2013; 32:964–977.

22. Coresh J, Turin TC, Matsushita K, et al. Decline in estimated glomerular filtration rate and subsequent risk of end-stage renal disease and mortality. *JAMA* 2014;311:2518–2531.

23. Lambers Heerspink HJ, Tighiouart H, Sang Y, et al. GFR decline

and subsequent risk of established kidney outcomes: a meta-analysis of 37 randomized controlled trials. *Am J Kidney Dis* 2014;64:860–866.

24. Herrington WG, Preiss D, Haynes R, et al. The potential for improving cardio-renal outcomes by sodium-glucose co-transporter-2 inhibition in people with chronic kidney disease: a rationale for the EMPA-KIDNEY study. *Clin Kidney J* 2018;11:749–761.

25. Mol PG, Maciulaitis R, Vetter T. GFR decline as an end point for clinical trials in CKD: a view from Europe. *Am J Kidney Dis* 2014;64:838–840.

26. Baigent C, Herrington WG, Coresh J, et al. Challenges in conducting clinical trials in nephrology: conclusions from a Kidney Disease-Improving Global Outcomes (KDIGO) Controversies Conference. *Kidney Int* 2017;92:297–305.

27. Levey AS, Inker LA. Assessment of glomerular filtration rate in health and disease: a state of the art review. *Clin Pharmacol Ther* 2017;102:405–419.

28. Murata K, Baumann NA, Saenger AK, Larson TS, Rule AD, Lieske JC. Relative performance of the MDRD and CKD-EPI equations for estimating glomerular filtration rate among patients with varied clinical presentations. *Clin J Am Soc Nephrol* 2011;6:1963–1972.

29. Selvin E, Juraschek SP, Eckfeldt J, Levey AS, Inker LA, Coresh J. Within-person variability in kidney measures. *Am J Kidney Dis* 2013;61:716–722.