

## Online-Only Supplementary Material

### Table of Contents

Table S1 .....	p. 2
Table S2 .....	p. 3
Figure S1 .....	p. 4

**Table S1. Relative importance of variables for predicting end-stage kidney disease (ESKD) using the random forest model and the Cox proportional hazard model (multiple imputation method)**

<b>Random forest model</b>			
<b>C-statistic 0.905</b>			
<b>Upper 12 variables</b>	<b>Relative importance</b>		
eGFR	0.098		
Serum albumin	0.024		
HbA1c	0.019		
Proteinuria, positive	0.017		
Serum uric acid	0.008		
Serum bilirubin	0.007		
Red blood cell count	0.005		
ARB use	0.004		
Age	0.002		
Body mass index	0.001		
Hypertension, positive	0.001		
Thrombocyte count	0.000		
<b>Cox proportional hazard model</b>			
<b>C-statistic 0.905</b>			
<b>Upper 12 variables</b>	<b>HR</b>	<b>95% CI</b>	<b>P value</b>
eGFR	0.245	0.195-0.308	<0.001
Serum bilirubin	0.557	0.417-0.744	<0.001
HbA1c	1.736	1.513-1.993	<0.001
Serum albumin	0.609	0.514-0.720	<0.001
Proteinuria, positive	1.624	1.291-2.043	<0.001
Age	0.707	0.591-0.846	<0.001
Serum uric acid	1.258	1.051-1.506	0.012
ARB use	1.228	1.013-1.490	0.037
GLP-1R agonist use	0.816	0.682-0.977	0.027
Gender, men	1.156	0.967-1.382	0.113
Hypertension, positive	0.883	0.672-1.161	0.374
Body mass index	0.887	0.735-1.071	0.213

\* P<0.05. ARB, angiotensin II receptor blocker; GLP-1R, glucagon-like peptide 1 receptor.

**Table S2. Characteristics of the subjects in the external validation cohort compared with those in the development cohort**

<b>Variables</b>	<b>Validation cohort</b>	<b>Development cohort</b>	<b>P value</b>
n	5153	2549	
Age, years, median (IQR)	48.0(42.0-54.0)	57.0 (47.0-63.0)	<0.001 <sup>a</sup>
Gender, male, n (%)	4622 (89.7)	1432 (56.2)	<0.001 <sup>b</sup>
HbA1c, %, median (IQR)	6.9 (6.4-7.9)	7.0 (6.5-8.1)	0.0831 <sup>a</sup>
mmol/mol, median (IQR)	50.4 (45.3-60.6)	51.4 (46.3-62.6)	0.0831 <sup>a</sup>
Serum albumin, mg/dL, mean (SD)	4.5 (0.4)	4.0 (0.5)	<0.001 <sup>c</sup>
Serum bilirubin, mg/dL, median (IQR)	0.7 (0.6-1.0)	0.7 (0.5-0.9)	<0.001 <sup>a</sup>
eGFR, mL min <sup>-1</sup> [1.73 m] <sup>-2</sup> , mean (SD)	80.6 (20.5)	81.8 (26.6)	0.0345 <sup>c</sup>
Proteinuria, n (%)			
Positive,	984 (19.1)	475 (29.7)	0.6492 <sup>b</sup>
Missing, n (%)	0	948 (37.2)	
Follow-up time, years,			
Median (IQR)	4.0 (3.9-4.6)	4.7 (2.3-8.8)	<0.001 <sup>a</sup>
Mean (SD)	4.3 (1.8)	5.6 (3.7)	<0.001 <sup>c</sup>
ESKD, n (%)	157 (3.1)	176 (6.2)	<0.001 <sup>c</sup>
Time to ESKD, years,			
Median (IQR)	3.1 (1.6-4.5)	2.5 (0.9-4.8)	<0.001 <sup>a</sup>
Mean (SD)	3.4 (2.2)	3.2 (2.9)	<0.001 <sup>c</sup>

a Calculated by Mann Whitney test, b Calculated by c2 test, c Calculated by t test. ESRD, end-stage kidney disease. Abbreviation: eGFR, estimated glomerular filtration rate; HbA1c, hemoglobin A1c.

