Chronic kidney disease

Thio, C. H. L.

DOI:
10.33612/diss.133648108

Evaluation of a genetic risk score based on creatinine-estimated glomerular filtration rate and its association with kidney outcomes

Chris HL Thio, Peter J van der Most, Ilja M Nolte, Pim van der Harst, Ute Bültmann, Ron T Gansevoort, Harold Snieder

Nephrol Dial Transplant. 2017;33(10):1757-64.

CHAPTER

5

ABSTRACT

Introduction. Cross-sectional GWAS on creatinine-estimated GFR ($eGFR_{crea}$) identified 53 SNPs. These SNP effects can be aggregated into a Genetic Risk Score (GRS) for chronic kidney disease (CKD). To assess its clinical utility, we examined associations with creatinine-estimated kidney outcomes, both cross-sectionally and longitudinally. Additionally, we examined associations with cystatin C-estimated kidney outcomes to verify that a GRS based on $eGFR_{crea}$ SNPs represents the genetics underlying kidney function.

Methods. In the community-based PREVEND Study, we assessed $eGFR_{crea}$ and $eGFR_{cysc}$ at baseline and four follow-up examinations. The GRS comprised 53 SNPs for $eGFR_{crea}$ weighted for reported effect-sizes. We adjusted for baseline demographics and renal risk factors.

Results. We included 3649 subjects (median age 49 years, 52% male, median follow-up 11 years, N=85 baseline CKD, N=154 incident CKD). At baseline, a higher GRS associated with lower $eGFR_{crea}$ (adjusted B (95%CI) = -2.05 (-2.45;-1.65) mL/min/1.73m^2, p<0.001) and higher CKD prevalence (adjusted OR (95%CI)= 1.41 (1.12;1.77), p=0.002). During follow-up, a higher GRS associated with higher CKD incidence (adjusted HR (95%CI)= 1.28 (1.09;1.50), p=0.004), but no longer significantly after adjustment for baseline eGFR. No significant association with $eGFR_{crea}$ decline was found. Associations with cystatin C-estimated outcomes were similar.

Conclusions. The GRS robustly associated with baseline CKD and eGFR, independent of known risk factors. Associations with incident CKD were likely due to low baseline eGFR, not accelerated eGFR decline. The GRS for $eGFR_{crea}$ likely represents the genetics underlying kidney function, not creatinine metabolism or underlying etiologies. To improve clinical utility of GWAS results for CKD, these need to specifically address eGFR decline and CKD incidence.

INTRODUCTION

Chronic kidney disease (CKD) is a heterogeneous group of diseases defined by the presence of sustained reduced kidney function or kidney damage. Strong evidence exists for a genetic component to CKD risk: CKD has been observed to aggregate in families[1-3] and heritability estimates are reported to range between 30 and 75%[4-9]. Furthermore, genome-wide association studies (GWAS) in populations of European ancestry have identified common genetic variants associated with CKD and kidney function markers[10-14] . The largest and most comprehensive genetic study is a cross-sectional meta-analysis of GWASs, in which single nucleotide polymorphisms (SNPs) at 53 loci were found to be associated with creatinine-estimated eGFR ($eGFR_{crea}$)[15].

The individual SNPs identified in this meta-analysis can be combined into a genetic risk score (GRS)[16-18], which summarizes individual genetic predisposition to CKD. Such a GRS is a potentially useful tool in etiological and predictive studies of CKD. However, because the SNPs were identified in a cross-sectional GWAS design, it is uncertain whether a GRS is associated with longitudinal outcomes. Furthermore, there is overlap between the 53 loci from the aforementioned meta-analysis and loci identified in a large GWAS on serum creatinine[11,12]. Therefore, it is difficult to discern whether a GRS corresponds to kidney function per se or partly reflects creatinine production/secretion.

The main study aim was to evaluate the applicability of a GRS, comprising 53 SNPs identified in cross-sectional GWAS on $eGFR_{crea}$, in longitudinal outcomes. To this end, we tested three hypotheses. First, we tested the hypothesis that the GRS would be associated with kidney outcomes, not only cross-sectionally (i.e. with baseline CKD, baseline eGFR), but also longitudinally (i.e. with incident CKD, eGFR decline). Second, to assess whether the GRS is a true representation of a genetic component to kidney function, we hypothesized that the GRS would also be associated with GFR estimates not based on serum creatinine. We therefore compared the associations of the GRS with $eGFR_{crea}$ to those of the GRS with an serum cystatin C-estimated GFR ($eGFR_{cysc}$)[19]. Third, to rule out that the GRS represents a component to kidney damage rather than kidney function, we hypothesized that the GRS would not be associated with albuminuria (i.e. urinary albumin excretion, UAE).

METHODS

Study population and design

We used data from the Prevention of REnal and Vascular ENdstage Disease (PREVEND) cohort study[20]. PREVEND was initiated to investigate the natural course of increased urinary albumin levels and its association to renal and vascular outcomes. Details of this study have been described elsewhere. In brief, 8592 individuals, sampled from the general population of Groningen, the Netherlands, underwent extensive examination between 1997-1998. The four follow-up examinations were completed in 2003, 2006, 2008, and 2012. Included were 3649 subjects of whom GWAS data were available. All subjects gave written informed consent. The PREVEND Study was approved by the medical ethics committee of the University Medical Center Groningen and conducted in accordance with the Helsinki Declaration guidelines.

Genetic risk scores

Genotyping details for PREVEND were described previously[21]. In brief, genotyping was performed on the Illumina CytoSNP12 v2 chip. Variants were imputed to 1000G[22], phase 1 version 3, using Minimac software[23]. Population stratification was assessed by principal component analysis; samples with Z-score>3 for any of the first five principal components were excluded, i.e. outlying individuals were removed because of likely divergent ancestry[24]. Samples with a call rate<95%, duplicates, and sex discrepancies were excluded. Markers with call rate>95%, Hardy-Weinberg equilibrium p-value≥$1x10^{-5}$, and minor allele frequency (MAF)≥1% were included. From the resulting GWAS data, we extracted the genotypes of the 53 SNPs that were identified in a recent meta-analysis of GWAS on $eGFR_{crea}$ in European populations[15]. Designated risk alleles were those associated with lower eGFR. Genotypes were represented as continuous allelic dosages from 0 to 2, reflecting an additive model [25]. A weighted GRS was defined as the sum of the risk alleles weighted for their published regression coefficient. Therefore, a higher GRS corresponds to higher susceptibility to impaired kidney function. For ease of interpretation, effects are reported per standard deviation (sd) higher GRS.

Outcome measurements and definition

At each examination, participants collected two consecutive 24h-urine specimens after thorough instruction. Participants were asked to avoid heavy exercise as

much as possible before urine collection, and instructed to postpone urine collection in case of urinary tract infection, menstruation, or fever. The collected urine was stored cold (4°C) for a maximum of four days before handing it in. After this, urine specimens were stored at -20°C. Fasting blood samples were obtained and stored at -80°C.

Measurement of serum creatinine was performed by an enzymatic method on a Roche Modular analyzer using reagents and calibrators from Roche (Roche Diagnostics, Mannheim, Germany), with intra- and interassay coefficients of variation of 0.9% and 2.9%, respectively. Serum cystatin C concentration was measured by a Gentian cystatin C Immunoassay (Gentian AS Moss, Norway) on a Modular analyzer (Roche Diagnostics). Cystatin C was calibrated directly using the standard supplied by the manufacturer (traceable to the International Federation of Clinical Chemistry Working Group for Standardization of Serum Cystatin C)[26] The intra- and interassay coefficients of variation were <4.1% and <3.3%, respectively. Urinary albumin concentration (UAC) was measured by nephelometry with a lower threshold of detection of 2.3mg/L, and intra- and interassay coefficient of variation of 2.2% and 2.6%, respectively (Dade Behring Diagnostic, Marburg, Germany). UAC was multiplied by urine volume to obtain a value of UAE in mg/24h. The two 24h-urinary albumin values of each subject per examination were averaged.

We calculated $eGFR_{crea}$ from serum creatinine and $eGFR_{cysc}$ from serum cystatin C, using the corresponding CKD-EPI equations [19]. We defined CKD_{crea} as $eGFR_{crea}$<60ml/min/1.73m^2, CKD_{cysc} as $eGFR_{cysc}$<60ml/min/1.73m^2, and CKD_{UAE} as UAE≥30mg/24h. Incident cases were those free of CKD at baseline who developed CKD during follow-up. In secondary analyses, we used the CKD-EPI equation for both serum creatinine and cystatin C to calculate $eGFR_{crea\text{-}cysc}$ [27]. Furthermore, a definition of CKD based on KDIGO guidelines (CKD_{KDIGO}, $eGFR_{crea\text{-}cysc}$<60ml/min/1.73m^2 and/or UAE≥30mg/24h) was used[28].

Covariates

We selected the following renal risk factors as covariates: age, sex, body-mass index (BMI, weight/height2 [kg/m^2]), current smoking (self-reported yes/no), diabetes (fasting glucose>7.0mmol/L, non-fasting glucose>11.0mmol/L, anti-diabetic treatment, or self-reported), hypertension (systolic blood pressure>140mmHg, diastolic blood pressure>90mmHg, blood pressure lowering treatment, or self-

reported), hypercholesterolemia (total cholesterol≥6.21mmol/L, lipid lowering treatment, or self-reported), and history of cardiovascular disease (CVD, any past cardio/cerebrovascular event or intervention). Covariates were collected at baseline by means of questionnaires, anthropometry, and pharmacy records.

Statistical analyses

Analyses were performed using R3.3.1 and SPSS23.0 (IBM Corporation). Two-sided significance level for analyses was set at α=0.05 unless stated otherwise.

Baseline characteristics

Baseline characteristics were examined for the total population. One-way ANOVA, Jonckheere-Terpstra, and χ^2-tests were used to examine linear trends of characteristics across tertiles of GRS. In subsequent analyses, GRS was treated as a continuous variable. We examined age and sex-adjusted associations of all 53 individual SNPs with baseline $eGFR_{crea}$ and $eGFR_{cysc}$ using ordinary least squares (OLS) regression.

Cross-sectional associations of the GRS with CKD prevalence and baseline eGFR

Logistic regression was used to examine the association of the continuous GRS with baseline CKD_{crea}. We adjusted for covariates by adding incremental groups of covariates in order to distinguish confounding effects of demographics and risk factors. Group 1 consisted of age and sex; group 2 additionally included BMI, smoking, diabetes, hypertension, hypercholesterolemia, and history of CVD.

We examined the association of the GRS with continuous $eGFR_{crea}$ using OLS regression. We adjusted for covariates as described above. Analyses were repeated for baseline $eGFR_{cysc}$ and prevalent CKD_{cysc}.

Longitudinal associations of the GRS with CKD incidence and eGFR decline

Cox regression models were used to examine the association of continuous GRS with incident CKD_{crea}. To estimate time to incident CKD_{crea}, we used a midpoint imputation technique. In this analysis, we corrected for baseline $eGFR_{crea}$ in addition to the previously listed renal risk factors. Subjects were censored at death or date of last visit.

Linear mixed-effects (LME) analysis was performed to examine the association of the GRS with eGFR decline. We modelled $eGFR_{crea}$ as a function of time since

baseline (per year). We specified a model with random intercept, random coefficient for time, and unstructured covariance matrix. The GRS, time, and covariates were included as fixed effects. A two-way interaction term between GRS and time was introduced to assess whether $eGFR_{crea}$ decline differed by values of the GRS. Analyses were repeated with the outcomes $eGFR_{cysc}$ decline and incident CKD_{cysc}.

Associations with UAE

We repeated the cross-sectional and longitudinal analyses described above to examine associations of a GRS with renal outcomes based on elevated UAE. Continuous UAE was transformed by its natural logarithm to approach normality (ln(UAE)), in OLS regression and LME analyses.

Secondary analyses

We repeated all analyses using $eGFR_{crea-cysc}$ and CKD_{KDIGO} as outcome. Furthermore, we constructed two alternative GRS. The first alternative GRS comprised 49 SNPs that were significant in the meta-analysis by Gorski et al.[14], with the second alternative comprising all 63 SNPs identified in either the Pattaro (53 SNPs) and the Gorski study (10 additional SNPs).

RESULTS

Baseline characteristics

Baseline characteristics of the 3649 subjects are presented in **Table 1**. In univariable analyses, a higher tertile for the GRS was associated with higher serum creatinine and cystatin C levels (p_{trend}<0.001); higher prevalence of CKD_{crea} (p_{trend} =0.002) and CKD_{cysc} (p_{trend} =0.01); lower $eGFR_{crea}$ (p_{trend} end<0.001) and lower $eGFR_{cysc}$ (p_{trend} <0.001); lower UAE (p_{trend} <0.001). No associations with CKD_{UAE} were found. We found no associations with age, sex, BMI, smoking status, diabetes, hypertension, hypercholesterolemia, or history of CVD.

Details of the 53 SNPs used in the calculation of the GRS and age- and sex- adjusted estimates of their association to baseline $eGFR_{crea}$, baseline $eGFR_{cysc}$, and ln(UAE) are listed in **Supplementary Table S1A**. Out of 53 SNPs, 22 reached nominal significance (one-sided p<0.05), while three were significant when a Bonferroni correction for 53 tests ($p<9.4x10^{-4}$) was applied.

Table 1. Baseline characteristics of the cohort stratified by tertiles of the Genetic Risk Score

	Total	GRS			P_{trend}
		low	medium	high	
N	3649	1216	1217	1216	n/a
Age, years	49 [39-60]	49 [40-60]	49 [39-59]	49 [39-60]	0.954
Males, %	52%	51%	51%	52%	0.598
BMI, kg/m²	26 (4.3)	26 (4.2)	26 (4.4)	26 (4.2)	0.816
BMI ≥30, %	16%	16%	16%	16%	0.868
Current smoker, %	35%	35%	35%	36%	0.420
Hypertension, %	34%	32%	36%	34%	0.521
SBP, mmHg	129 (20)	129 (20)	129 (20)	129 (20)	0.612
DBP, mmHg	74 (9.9)	74 (9.9)	74 (10)	74 (9.9)	0.887
BP lowering medication, %	12%	13%	14%	11%	0.658
Diabetes, %	3.9%	3.7%	3.4%	4.7%	0.210
Glucose, mmol/L	4.7 [4.3-5.1]	4.7 [4.4-5.1]	4.7 [4.3-5.2]	4.7 [4.4-5.1]	0.926
Anti-diabetic medication, %	1.3%	1.2%	1.0%	1.8%	0.843
Hypercholesterolemia, %	31%	31%	32%	31%	1.000
Total cholesterol, mmol/L	5.7 (1.1)	5.7 (1.1)	5.6 (1.1)	5.7 (1.1)	0.744
Lipid lowering medication, %	3.6%	4.8%	3.7%	2.7%	0.499
History of CVD, %	4.2%	3.9%	5.1%	3.8%	0.920
Serum creatinine, mg/dL	0.82 (0.18)	0.79 (0.16)	0.82 (0.18)	0.85 (0.19)	<0.001
$eGFR_{crea}$, mL/min/1.73m²	96 (16)	98 (15)	96 (16)	94 (16)	<0.001
CKD_{crea}: $eGFR_{crea}$ <60, %	2.5%	1.4%	2.7%	3.4%	0.002
Serum cystatin C, mg/L	0.90 (0.18)	0.88 (0.17)	0.90 (0.19)	0.92 (0.18)	<0.001
$eGFR_{cysc}$, mL/min/1.73m²	92 (19)	94 (19)	92 (19)	90 (19)	<0.001
CKD_{cysc}: $eGFR_{cysc}$ <60, %	5.9%	4.9%	5.3%	7.4%	0.010
$eGFR_{crea\text{-}cysc}$	94 (17)	97 (17)	95 (17)	92 (17)	<0.001
CKD_{KDIGO} :$eGFR_{crea\text{-}cysc}$ <60 or UAE≥30 , %	20%	21%	20%	19%	0.297
UAE, mg/24h	10.6 [6.6-21]	11.5 [7.0-23]	10.3 [6.6-20]	10.2 [6.4-20]	<0.001
UAE ≥30 , %	17%	19%	17%	17%	0.172
No of risk alleles	57 (4.5)	52 (2.7)	57 (1.7)	62 (2.5)	<0.001

Baseline characteristics of the cohort. Data is presented as mean (standard deviation), median [interquartile range], and percentage where appropriate. P-values for linear trend were calculated using one-way ANOVA, Jonckheere-Terpstra-tests, and χ^2-tests where appropriate.

Abbreviations: BMI, body mass index; CVD, cardiovascular disease; CKD, chronic kidney disease; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; GRS, genetic risk score; SBP, systolic blood pressure; UAE, urinary albumin excretion

Supplementary Figure S2 presents a plot of age- and sex-adjusted regression coefficients. These coefficients were obtained by OLS regression of individual SNPs on either $eGFR_{crea}$ and $eGFR_{cysc}$. Correlation between the regression coefficients on $eGFR_{crea}$ and $eGFR_{cysc}$ was moderate (Pearson r=0.51, p<0.001). The total least squares regression line showed fair agreement with the line of identity.

Cross-sectional associations of the GRS with baseline eGFR and CKD prevalence

We present cross-sectional results in **Table 2**. Per sd higher GRS, the odds of having CKD_{crea} at baseline increased by 41% (fully adjusted odds ratio (OR) (95%CI)=1.41 (1.12;1.77), p=0.002. A higher GRS was associated with lower $eGFR_{crea}$ (fully adjusted unstandardized coefficient B (95%CI)= -2.05 (-2.45;-1.65) mL/min/1.73m², p<0.001), independent of known risk factors. Effect sizes of the associations with CKD_{cysc} (adjusted OR (95%CI)= 1.27 (1.08;1.50), p=0.004) and with $eGFR_{cysc}$ (adjusted B (95%CI)= -1.63 (-2.11;-1.14) mL/min/1.73m², p<0.001) were smaller but showed a similar trend compared to those for creatinine-estimated outcomes. Estimates of the effect sizes of the GRS on both $eGFR_{crea}$ and $eGFR_{cysc}$ remained stable during incremental covariate adjustment.

Longitudinal associations of the GRS with eGFR decline and CKD incidence

We present longitudinal results in **Table 3**. A higher GRS was associated with higher incidence of CKD_{crea} after adjustment for known renal risk factors (adjusted hazard ratio (HR) (95%CI)=1.28 (1.09;1.50), p=0.003), but significance disappeared after additional adjustment for baseline $eGFR_{crea}$ (fully adjusted HR (95%CI)=1.05 (0.89;1.24), p=0.537). A higher GRS was not associated with steeper decline of $eGFR_{crea}$ (fully adjusted B (95%CI)= -0.01 (-0.04;0.03) mL/min/1.73m² per year, p=0.655). Inclusion of interaction terms between baseline renal risk factors and time did not change estimates of the effects between the GRS and eGFR decline (data not shown).

Similar associations were found with $eGFR_{cysc}$ decline (fully adjusted B (95%CI)= -0.03 (-0.07;0.01) mL/min/1.73m² per year, p=0.167) and incident CKD_{cysc} (adjusted HR (95%CI)=1.17 (1.03;1.32), p=0.014. The association with incident CKD_{cysc} lost significance after additional adjustment for baseline $eGFR_{cysc}$ (fully adjusted HR (95%CI)=1.06 (0.94;1.20), p=0.336).

Table 2. Cross-sectional associations of the Genetic Risk Score with selected kidney outcomes at baseline

	Dichotomous outcomes					Continuous outcomes			
	Prevalent CKD_{crea} (85 cases / N= 3397)	**Prevalent CKD_{cysc}** (199 cases / N=3394)	**Prevalent CKD_{UAE}** (635 cases / N=3614)	**Prevalent CKD_{KDIGO}** (684 cases / N=3423)		**$eGFR_{crea}$** (N=3397)	**$eGFR_{cysc}$** (N=3394)	**$eGFR_{crea-cysc}$** (N=3394)	**ln(UAE)** (N=3614)
	OR (95%CI)	**OR (95%CI)**	**OR (95%CI)**	**OR (95%CI)**		**B (95%CI)**	**B (95%CI)**	**B (95%CI)**	**B (95%CI)**
Model 1	1.38 (1.11;1.71)**	1.21 (1.05;1.40)**	0.92 (0.85;1.00)	0.93 (0.86;1.01)		-1.87 (-2.39; -1.34)***	-1.45 (-2.10;-0.81)***	-1.85 (-2.42;-1.28)***	-0.05 (-0.08;-0.01)**
Model 2	1.41 (1.13;1.76)**	1.27 (1.08;1.48)**	0.93 (0.85;1.01)	0.94 (0.86;1.02)		-2.04 (-2.44;-1.64)***	-1.66 (-2.15;-1.16)***	-2.04 (-2.46;-1.61)***	-0.04 (-0.07;-0.01)**
Model 3	1.41 (1.12;1.77)**	1.27 (1.08;1.50)**	0.92 (0.84;1.01)	0.93 (0.85;1.02)		-2.05 (-2.45;-1.65)***	-1.63 (-2.11;-1.14)***	-2.02 (-2.45;-1.60)***	-0.04 (-0.07;-0.01)**

Estimates from linear and logistic regression analyses. Data is presented as regression coefficient B (95% confidence interval), or odds ratio OR (95% confidence interval), per standard deviation (sd) of GRS. Definitions and abbreviations: eGFR, estimated glomerular filtration rate (mL/min/1.73m²) ; ln(UAE), natural logarithm (ln) of urinary albumin excretion (ln mg/24h) ; $CKD_{crea/cysc}$, chronic kidney disease ($eGFR_{crea/cysc}$ <60mL/min/1.73m²) ; CKD_{UAE} (UAE ≥30 mg/24h) ; CKD_{KDIGO} , ($eGFR_{crea-cysc}$ <60mL/min/1.73m² and/or UAE ≥30 mg/24h) ; GRS, genetic risk score. *p<0.05, **p<0.01, ***p<0.001.

Model 1: GRS
Model 2: model 1 + age + sex,
Model 3: model 2 + BMI + smoking + diabetes + hypertension + hypercholesterolemia + history of cardiovascular disease.

Table 3. Longitudinal associations of the Genetic Risk Score with selected kidney outcomes during follow-up

	Dichotomous outcomes					Continuous outcomes			
	Incident CKD_{crea} (154 cases / N= 2731)	**Incident CKD_{cysc}** (279 cases / N=2659)	**Incident CKD_{UAE}** (368 cases / N=2493)	**Incident CKD_{KDIGO}** (411 cases / N=2296)		**Δ$eGFR_{crea}$** (N=3447)	**Δ$eGFR_{cysc}$** (N=3447)	**Δ$eGFR_{crea-cysc}$** (N=3447)	**Δln(UAE)** (N=3619)
	HR (95%CI)	**HR (95%CI)**	**HR (95%CI)**	**HR (95%CI)**		**B (95%CI)**	**B (95%CI)**	**B (95%CI)**	**B (95%CI)**
Model 1	1.19 (1.02;1.40)*	1.08 (0.96;1.21)	0.94 (0.85;1.04)	1.02 (0.93;1.13)		-0.01 (-0.04;0.03)	-0.03 (-0.07;0.01)	-0.02 (-0.05;0.02)	0.001 (-0.001;0.004)
Model 2	1.28 (1.09;1.50)**	1.15 (1.02;1.29)*	0.96 (0.87;1.06)	1.07 (0.97;1.18)		-0.01 (-0.04;0.03)	-0.03 (-0.07;0.01)	-0.02 (-0.05;0.02)	0.001 (-0.001;0.004)
Model 3	1.28 (1.09;1.50)**	1.17 (1.03;1.32)*	0.95 (0.86;1.06)	1.06 (0.96;1.17)		-0.01 (-0.04;0.03)	-0.03 (-0.07;0.01)	-0.02 (-0.05;0.02)	0.001 (-0.001;0.004)
Model 4	1.05 (0.89;1.24)	1.06 (0.94;1.20)	1.03 (0.93;1.14)	1.11 (1.00;1.22)^		-	-	-	-

Estimates from Cox regression and LME analyses. Data is presented as hazard ratio HR (95% confidence interval), or regression coefficient B (95% confidence interval), per standard deviation of GRS. Definitions and abbreviations: Δ eGFR, annual change in estimated glomerular filtration rate (mL/min/1.73m² per year) ; Δ ln(UAE), annual change in natural logarithm (ln) of urinary albumin excretion (ln (mg/24h) per year) ; $CKD_{crea/cysc}$, chronic kidney disease ($eGFR_{crea/cysc}$ <60mL/min/1.73m²) ; CKD_{UAE} (UAE ≥30 mg/24h) ; CKD_{KDIGO} , ($eGFR_{crea-cysc}$ <60mL/min/1.73m² and/or UAE ≥30 mg/24h); GRS, genetic risk score. *p<0.05, **p<0.01, ***p<0.001.

Model 1: GRS
Model 2: model 1 + age + sex
Model 3: model 2 + BMI + smoking + diabetes + hypertension + hypercholesterolemia + history of cardiovascular disease
Model 4: model 3 + baseline eGFR (adjusted for baseline UAE instead of eGFR) (^adjusted for both baseline eGFR and UAE)

Association of the GRS with UAE

Results of analyses on UAE are presented in **Table 2-3**. A higher GRS was associated with lower ln(UAE) (fully adjusted B (95%CI)= -0.04 (-0.07;-0.01) ln(mg/24h), p=0.004) but not with higher prevalence of CKD_{UAE} (fully adjusted OR (95%CI)=0.92 (0.84;1.01), p=0.074). No longitudinal associations of GRS with kidney damage were observed: a higher GRS was neither associated with steeper increase of ln(UAE) (fully adjusted B (95%CI)=0.001 (-0.001;0.004) ln(mg/24h) per year, p=0.297) nor with higher incidence of CKD_{UAE} (fully adjusted HR (95%CI)=1.03 (0.93;1.14), p=0.360).

Analyses with 24h-urinary albumin-to-creatinine ratio as outcome yielded similar results (data not shown).

Secondary analyses

Associations of the GRS with $eGFR_{crea-cysc}$ were consistent with those of the GRS with $eGFR_{crea}$ and $eGFR_{cysc}$. We found no cross-sectional or longitudinal association of the GRS with CKD_{KDIGO} (Table 2-3). Two alternative GRS, based on 49 SNPs ($GRS_{1000G-49}$) and 63 SNPs ($GRS_{1000G-63}$), were evaluated. Individual SNP-effects of these GRS are listed in **Supplementary Table S1B**. The GRSs showed similar but slightly weaker associations compared to our main GRS (**Supplementary Table S3-7**).

DISCUSSION

In this population based, longitudinal cohort study, we evaluated the effects of a GRS comprising 53 $eGFR_{crea}$-SNPs on kidney outcomes. To this end, we tested cross-sectional and longitudinal associations of this GRS with CKD_{crea} and $eGFR_{crea}$ and compared these associations to those with CKD_{cysc} and $eGFR_{cysc}$. Cross-sectional associations of the GRS with the kidney outcomes, CKD_{crea} and $eGFR_{crea}$, were modest but robust, corroborating the literature. In longitudinal analyses, we observed no associations with kidney function decline. The GRS was associated with incidence of CKD_{crea}, but this was likely due to lower baseline eGFR rather than accelerated kidney function decline. In comparison to associations with $eGFR_{crea}$, associations with $eGFR_{cysc}$ were smaller but showed a similar trend. Higher GRS was not associated with kidney damage markers. Furthermore, all associations of the GRS with kidney outcomes were independent of renal risk factors. These data suggest that the GRS is a true representation of the genetics underlying kidney function, as opposed to creatinine metabolism, kidney damage, or related etiologies such as hypertension/diabetes.

In secondary analyses, we confirmed associations with $eGFR_{crea-cysc}$, currently the best estimate for kidney function for large population-based studies [19,29]. We found no association of the GRS with CKD_{KDIGO} as outcome. This is likely due to the fact that this GRS was optimized for eGFR as outcome and not urinary albumin; in our sample, CKD_{KDIGO} was predominantly characterized by elevated urinary albumin rather than diminished kidney function. Two alternative GRS ($GRS_{1000G-49}$ and $GRS_{1000G-63}$), yielded similar results but proved to be slightly less powerful predictors of kidney function and CKD in this sample.

Previously, two similar GRSs based on $eGFR_{crea}$ SNPs were investigated in ~2500 participants with ~11 years of follow-up from the Framingham Heart Study. O'Seaghdha et al. calculated a 16-SNP GRS for $eGFR_{crea}$ [17]. This sample of the Framingham cohort was revisited by Ma et al.[18], who updated the GRS with 37 additional SNPs, that is the same 53 as the present study. Both of these GRS were independently associated with incident CKD ($eGFR_{crea}$<60mL/min/1.73m^2), although neither of these GRSs improved prediction and/or discrimination beyond clinical risk factors (age, sex, BMI, eGFR, hypertension, diabetes, proteinuria). Interestingly, they reported associations of a higher GRS with a higher incidence of CKD to be independent of baseline eGFR, hence an accelerated deterioration of kidney function in those with a higher GRS. Such an effect was also suggested by Böger et al.[30] in a study of eGFR related loci identified by GWAS. In 26,308 individuals of European ancestry, the associations of 16 separate SNPs known at the time with incident CKD were examined. Of these 16 SNPs, six (mapping to *UMOD*, *PRKAG2*, *LASS2*, *DAB2*, *DACH1*, and *STC1*) were significantly ($p<0.05$) associated with incident CKD (eGFR<60mL/min/1.73m^2), even after correction for baseline eGFR. Similar to the findings of O'Seaghdha and Ma et al, this implies that several SNPs associate with eGFR decline. In contrast, in the present study we could not corroborate such an effect on CKD incidence or eGFR decline: the association of GRS with incident CKD was not significant after adjustment for baseline eGFR, and there was no significant association between the GRS and eGFR decline.

A possible explanation for this discrepancy is the potential overestimation of the effect of the GRS by O'Seaghdha and Ma et al. due to the participation of the Framingham Cohort Study in the discovery phase of the meta-analysis[12,15]. Similarly, overestimation of individual SNP effects may have occurred in the

study by Böger et al, given that seven of the eight cohorts participating in that study were part of the discovery GWAS[12]. Such overlap in discovery and validation cohorts might result in inflated effect sizes[31]. The PREVEND study was not part of the original discovery GWAS, ensuring its independence and suitability as a validation cohort for evaluation of a GRS based on $eGFR_{crea}$ SNPs. This potential overestimation possibly also explains that in our study, the GRS explained only 1.66% of variance of baseline $eGFR_{crea}$, whereas in the original GWAS, the explained variance of $eGFR_{crea}$ by the combined loci was 3.22%[15].

Notwithstanding these discrepancies, the combined data suggest that the genetics underlying kidney function are, at least partly, distinct from that underlying kidney function decline and/or kidney disease susceptibility. Our results indicate that a GRS based on cross-sectional GWAS results on kidney function is not clinically applicable (e.g. in the prediction of CKD risk). A GRS would be more applicable if SNPs associated with kidney function decline and/or CKD incidence were used, as these would likely better represent disease susceptibility. Unfortunately, there is paucity of data on genetic loci associated with kidney function decline or CKD incidence. To the best of our knowledge, only one study by Gorski et al. performed a GWAS for kidney function decline phenotypes[32]. In this study, only one SNP mapping to *UMOD* (which was also implicated in prior GWAS on cross-sectional $eGFR_{crea}$) was significantly associated with eGFR change in the general population, while two novel loci, *CDH23* and *GALNT15/GALNT11* were only suggestively associated with eGFR change in CKD patients, and rapid decline in the general population, respectively. To benefit clinical applicability, we argue that future GWAS should focus on disease susceptibility genes, i.e. loci associated with eGFR decline and/or CKD incidence. We found a higher GRS to be associated with lower UAE, i.e. lower risk of kidney damage, which is surprising for two reasons. First, a prior family study, using bivariate variance component linkage analysis techniques, found a low genetic correlation between eGFR and UACR (r_g=0.002 in African Americans, not reported for European Americans)[5]. Second, there is no overlap in genome-wide significant markers for eGFR and albuminuria in the general population[33, 34]. Due to this apparent lack of genetic overlap, it is believed that eGFR and albuminuria have distinct genetic underpinnings. To our knowledge, we are the first to observe this counterintuitive association with the updated 53 SNP GRS. Although the correlation between the GRS and ln(UAE) was weak (r=-0.043), it is unlikely to be a chance finding: in an earlier study by Ellis et al.

a weighted GRS (comprising 16 SNPs associated with eGFRcrea) was associated with both lower eGFR and with lower UACR[35]. The authors attributed this effect to the A-allele of rs17319721, a SNP mapping to *SHROOM3*, because exclusion of this SNP from their GRS attenuated the effect on UACR. In previous GWAS, the *SHROOM3* SNP was found to be associated with $eGFR_{crea}$[12], and suggestively with UACR ($p=7.0x10^{-7}$)[15,33,34]. In the present study, exclusion of this SNP from the updated GRS did not attenuate the effect (data not shown). Therefore, it is possible that, in addition to *SHROOM3*, other loci discovered in the recent meta-analysis on $eGFR_{crea}$ might have pleiotropic effects on both eGFR and albuminuria. We therefore performed a query in LDHub v1.3.1, a platform for LD-score regression which uses original GWAS summary statistics[36, 37]. LD Hub showed a modest genetic correlation between $eGFR_{crea}$ and UACR ($r_g=0.388$, $p<0.001$), and a suggestive genetic correlation between $eGFR_{cysc}$ and UACR ($r_g=0.195$, $p=0.087$), in the same direction as our findings (i.e. higher eGFR~higher UACR). These correlations suggest that there is at least partial overlap in the genetics underlying eGFR and albuminuria. Addressing the question of pleiotropy is beyond the scope of the present study and requires dedicated analysis in larger samples.

A number of SNPs identified in the GWAS on $eGFR_{crea}$ may be linked to loci related to creatinine production or secretion, hence not with kidney function per se[38]. We therefore examined two SNPs mapping to loci known to be related to creatinine metabolism: rs2467853 which maps to the creatinine production locus *GATM*[39] and rs316009 which maps to the creatinine secretion locus SLC22A2[40]. For both SNPs, we observed an inconsistency in the direction of effect for baseline $eGFR_{crea}$ and $eGFR_{cysc}$ (see **Supplementary Table 1A**), suggesting that these loci are indeed not related to kidney function. Exclusion of these SNPs led to a slightly improved GRS: effects of this GRS on $eGFR_{crea}$ and $eGFR_{cysc}$ more closely resembled each other than those of the main GRS, although this improvement was only slight (data not shown). Our conclusions therefore remain unchanged. Future, functional studies may investigate other presumptive creatinine-related loci. The exclusion of such loci may result in a GRS that more accurately reflects genetic predisposition to kidney function.

To our knowledge, we are the first study that examined the association between a GRS comprising 53 SNPs and eGFR decline. Strengths of this study include the availability of serially measured creatinine and cystatin C, as well as two 24h-urinary

albumin at each examination, during a considerable follow-up duration of 11 years. A major strength of PREVEND is its independence from the discovery GWAS that identified the 53 SNPs used in the GRS, resulting in unbiased effect estimates of the GRS. Given that participants of the PREVEND GWAS sample are of European ancestry, we cannot generalize to other ethnicities. Finally, we could not calculate genetic correlations between eGFR levels and eGFR decline as GWAS summary results for eGFR decline were currently not available.

In conclusion, a GRS comprising 53 SNPs showed modest but robust associations with cross-sectional CKD outcomes based on $eGFR_{crea}$. These associations were confirmed with $eGFR_{cysc}$, which highlights the potential usefulness of a GRS as a representation of the genetics underlying kidney function. However, no longitudinal associations with incident CKD or eGFR decline were found. Given these results, we question the clinical utility of cross-sectional GWAS results on kidney function. We suggest that future GWAS specifically examine genetic associations with eGFR decline and/or CKD incidence. These GWAS may identify loci that, when incorporated into a GRS, will improve the clinical utility of this score, e.g. in predicting onset of CKD.

CONFLICT OF INTEREST STATEMENT

None of the authors declare a conflict of interests.

ACKNOWLEDGEMENTS

The PREVEND Study was funded by grants from the Dutch Kidney Foundation

All Supplementary material can be accessed via the following link:
www.academic.oup.com/ndt/article/33/10/1757/4774601

REFERENCES

1. Skrunes R, Svarstad E, Reisaeter AV, Vikse BE. Familial clustering of ESRD in the norwegian population. *Clin J Am Soc Nephrol*. 2014;9(10):1692-1700.

2. Satko SG, Sedor JR, Iyengar SK, Freedman BI. Familial clustering of chronic kidney disease. . 2007;20(3):229-236.

3. Lei HH, Perneger TV, Klag MJ, Whelton PK, Coresh J. Familial aggregation of renal disease in a population-based case-control study. *J Am Soc Nephrol*. 1998;9(7):1270-1276.

4. Arpegard J, Viktorin A, Chang Z, de Faire U, Magnusson PK, Svensson P. Comparison of heritability of cystatin C- and creatinine-based estimates of kidney function and their relation to heritability of cardiovascular disease. *J Am Heart Assoc*. 2015;4(1):e001467.

5. Leon JM, Freedman BI, Miller MB, et al. Genome scan of glomerular filtration rate and albuminuria: The HyperGEN study. *Nephrol Dial Transplant*. 2007;22(3):763-771.

6. Fox CS, Yang Q, Cupples LA, et al. Genomewide linkage analysis to serum creatinine, GFR, and creatinine clearance in a community-based population: The framingham heart study. *J Am Soc Nephrol*. 2004;15(9):2457-2461.

7. Hunter DJ, Lange M, Snieder H, et al. Genetic contribution to renal function and electrolyte balance: A twin study. *Clin Sci (Lond)*. 2002;103(3):259-265.

8. Langefeld CD, Beck SR, Bowden DW, Rich SS, Wagenknecht LE, Freedman BI. Heritability of GFR and albuminuria in caucasians with type 2 diabetes mellitus. *Am J Kidney Dis*. 2004;43(5):796-800.

9. Bochud M, Elston RC, Maillard M, et al. Heritability of renal function in hypertensive families of african descent in the seychelles (indian ocean). *Kidney Int*. 2005;67(1):61-69.

10. Kottgen A, Glazer NL, Dehghan A, et al. Multiple loci associated with indices of renal function and chronic kidney disease. *Nat Genet*. 2009;41(6):712-717.

11. Chambers JC, Zhang W, Lord GM, et al. Genetic loci influencing kidney function and chronic kidney disease. *Nat Genet*. 2010;42(5):373-375.

12. Kottgen A, Pattaro C, Boger CA, et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet*. 2010;42(5):376-384.

13. Gudbjartsson DF, Holm H, Indridason OS, et al. Association of variants at UMOD with chronic kidney disease and kidney stones—role of age and comorbid diseases. *PLoS Genet*. 2010;6(7):e1001039.

14. Gorski M, van der Most PJ, Teumer A, et al. 1000 genomes-based meta-analysis identifies 10 novel loci for kidney function. *Scientific Reports*. 2017;7:45040.

15. Pattaro C, Teumer A, Gorski M, et al. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat Commun*. 2016;7.

16. Morrison AC, Bare LA, Chambless LE, et al. Prediction of coronary heart disease risk using a genetic risk score: The atherosclerosis risk in communities study. *Am J Epidemiol*. 2007;166(1):28-35.

17. O'Seaghdha CM, Yang Q, Wu H, Hwang SJ, Fox CS. Performance of a genetic risk score for CKD stage 3 in the general population. *Am J Kidney Dis*. 2012;59(1):19-24.

18. Ma J, Yang Q, Hwang S, Fox CS, Chu AY. Genetic risk score and risk of stage 3 chronic kidney disease. *BMC nephrology*. 2017;18(1):32.

19. Inker LA, Schmid CH, Tighiouart H, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. *N Engl J Med*. 2012;367(1):20-29.

20. Pinto-Sietsma SJ, Janssen WM, Hillege HL, Navis G, De Zeeuw D, De Jong PE. Urinary albumin excretion is associated with renal functional abnormalities in a nondiabetic population. *J Am Soc Nephrol*. 2000;11(10):1882-1888.

21. Verweij N, Mateo Leach I, Isaacs A, et al. Twenty-eight genetic loci associated with ST-T-wave amplitudes of the electrocardiogram. *Hum Mol Genet*. 2016;25(10):2093-2103.

22. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467(7319):1061-1073.

23. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet*. 2012;44(8):955-959.

24. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. *Nat Protoc*. 2010;5(9):1564-1573.

25. Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. *Nature protocols*. 2011;6(2):121-133.

26. Grubb A, Blirup-Jensen S, Lindstrom V, et al. First certified reference material for cystatin C in human serum ERM-DA471/IFCC. *Clin Chem Lab Med*. 2010;48(11):1619-1621.

27. Stevens LA, Coresh J, Schmid CH, et al. Estimating GFR using serum cystatin C alone and in combination with serum creatinine: A pooled analysis of 3,418 individuals with CKD. *Am J Kidney Dis*. 2008;51(3):395-406.

28. Levey AS, de Jong PE, Coresh J, et al. The definition, classification, and prognosis of chronic kidney disease: A KDIGO controversies conference report. *Kidney Int*. 2011;80(1):17-28.

29. Inker LA, Levey AS, Tighiouart H, et al. Performance of glomerular filtration rate estimating equations in a community-based sample of blacks and whites: The multiethnic study of atherosclerosis. *Nephrology Dialysis Transplantation*. 2017:gfx042.

30. Böger CA, Gorski M, Li M, et al. Association of eGFR-related loci identified by GWAS with incident CKD and ESRD. *PLoS Genet*. 2011;7(9):e1002292.

31. Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. Pitfalls of predicting complex traits from SNPs. *Nature Reviews Genetics*. 2013;14(7):507-515.

32. Gorski M, Tin A, Garnaas M, et al. Genome-wide association study of kidney function decline in individuals of european descent. *Kidney Int*. 2015;87(5):1017-1029.

33. Boger CA, Chen MH, Tin A, et al. CUBN is a gene locus for albuminuria. *J Am Soc Nephrol*. 2011;22(3):555-570.

34. Teumer A, Tin A, Sorice R, et al. Genome-wide association studies identify genetic loci associated with albuminuria in diabetes. *Diabetes*. 2016;65(3):803-817.

35. Ellis JW, Chen MH, Foster MC, et al. Validated SNPs for eGFR and their associations with albuminuria. *Hum Mol Genet*. 2012;21(14):3293-3298.

36. Zheng J, Erzurumluoglu AM, Elsworth BL, et al. LD hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*. 2017;33(2):272-279.

37. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015.

38. Kottgen A. Genome-wide association studies in nephrology research. *Am J Kidney Dis*. 2010;56(4):743-758.

39. Wyss M, Kaddurah-Daouk R. Creatine and creatinine metabolism. *Physiol Rev*. 2000;80(3):1107-1213.

40. Reznichenko A, Snieder H, van den Born J, et al. CUBN as a novel locus for end-stage renal disease: Insights from renal transplantation. *PLoS One*. 2012;7(5):e36512.