#### **ORIGINAL ARTICLE**



# Polygenic risk scores for eGFR are associated with age at kidney failure

Kane E. Collins · Edmund Gilbert · Vincent Mauduit · Pukhraj Gaheer · Elhussein A. E. Elhassan · Katherine A. Benson, et al. [full author details at the end of the article]

Received: 2 September 2024 / Accepted: 2 January 2025 / Published online: 3 March 2025 © The Author(s) 2025

# Abstract

**Background** The genetic architecture of chronic kidney disease (CKD) is complex, including monogenic and polygenic contributions. CKD progression to kidney failure is influenced by factors including male sex, baseline estimated glomerular filtration rate (eGFR), hypertension, diabetes, proteinuria, and the underlying kidney disease. These traits all have strong genetic components, which can be partially quantified using polygenic risk scores. This paper examines the association between polygenic risk scores for CKD-related traits and age at kidney failure development.

**Methods** Genome-wide genotype data from 10,586 patients with kidney failure were compiled from 12 cohorts. Polygenic risk scores for hypertension, albuminuria, rapid decline in eGFR, decreased total kidney volume, and decreased eGFR were calculated using weights from published independent population-scale genome-wide association studies. The association between each polygenic risk score and age at kidney failure was investigated using logistic regression models. The association between polygenic risk score and age at kidney failure was also investigated separately for each primary kidney disease. **Results** Individuals in the highest 10% of polygenic risk score for decreased eGFR developed kidney failure 2 years earlier than those in the bottom 90% (49.9 years and 47.9 years, P = 5e-5). A standard deviation increase in decreased eGFR polygenic risk score was associated with increased odds of developing kidney failure before the age of 60 years (Odds ratio (OR) = 1.05; 95% CI 1.01–1.10; P = 0.01), as was high decreased eGFR polygenic risk score (OR = 1.26; 95% CI 1.08–1.46; P = 0.003).

**Conclusions** We conclude that decreased eGFR polygenic risk score explains a portion of the variation in age at development of kidney failure.

Gianpiero L. Cavalleri and Peter J. Conlon these authors contributed equally.

#### **Graphical abstract**



Keywords Polygenic burden · Age at onset · Kidney failure · Kidney function · Polygenic risk score

# Introduction

Chronic kidney disease (CKD), a condition characterized by reduced estimated glomerular filtration rate (eGFR) and/ or proteinuria, is a significant cause of morbidity and mortality on a global scale [1]. The irreversible loss of GFR which CKD causes, irrespective of the primary cause of kidney damage, often leads to kidney failure defined as eGFR below 15 mL/min/1.73 m<sup>2</sup> or a need for kidney replacement therapy. Kidney failure is a severe, life-limiting condition, where kidney function is critically impaired, and patients require dialysis or transplantation to stay alive. It affects over 800,000 individuals in the United States alone [2], resulting in enormous suffering, premature death, and financial cost. Risk factors, such as male sex, hypertension, diabetes, proteinuria, family history of CKD, lower eGFR, and genetic factors contribute to the progression of CKD [3].

Different causes of CKD can lead to kidney failure at different ages. Factors which have been associated with age at kidney failure include degree of proteinuria, total kidney volume, baseline eGFR and eGFR change [4]. In addition, therapeutic strategies exist to delay the age at which kidney failure develops. These focus on control of hypertension and blood glucose, through diet and lifestyle factors (e.g. exercise), or medications (e.g. ACE and SGLT2 inhibitors, or specific immunosuppressive medications for certain kidney conditions such as glomerulonephritis) [5, 6].

Although inherited monogenic forms of kidney diseases are estimated to account for as much as 70% of the overall prevalence of kidney failure in children and 10–15% in adults [7], the complex polygenic architecture of CKD, which interacts with environmental factors, influences disease progression [8]. Estimates of the heritability of CKD range between 36 and 75% [9].

A recent study with more than one million individuals has enabled the discovery of over 250 loci significantly associated with CKD [10]. Further, well-powered genome-wide association studies (GWAS) have been conducted for diverse traits related to incident CKD and disease progression to kidney failure, including hypertension [11], eGFR [10], albuminuria [12], total kidney volume [13], and rapid decline in eGFR [14]. Summary statistics from genome-wide association studies can be used to calculate polygenic risk scores (PRSs). These scores estimate the cumulative impact of common genetic variations on an individual's disease status, weighted by the estimated effect size [15]. For instance, individuals in the top 2% of polygenic risk score for eGFR have been reported to be at a threefold greater risk of developing CKD, relative to those in the bottom 98% of polygenic risk score [16]. In a cohort of more than 11,000 participants over the age of 70, it was shown that a standard deviation increase in eGFR polygenic risk score was associated with a decrease of 3.9 mL/min/1.73 m<sup>2</sup> [17].

Moreover, the polygenic burden for CKD has been reported to alter the penetrance of monogenic kidney disease. Those with autosomal dominant polycystic kidney disease (ADPKD) in the top tertile of polygenic burden for CKD have a significantly higher CKD risk than those in the lowest tertile of polygenic burden [18].

eGFR polygenic risk score has previously been associated with incident CKD stage 3 development [18, 19]. However, testing the association of eGFR polygenic risk score with kidney failure events has been limited to less than a few hundred events [19]. Moreover, the association of polygenic risk score for kidney traits including hypertension, albuminuria, and total kidney volume and age at kidney failure has not been evaluated. Using data from 10,586 patients with kidney failure, we hypothesize that the polygenic burden for CKD progression is associated with age at kidney failure.

# Materials and methods

#### Study design and inclusion criteria

To investigate the impact of polygenic burden for traits related to kidney function on age at kidney failure, we assembled a large cohort of patients with kidney failure (n = 10,586).

The inclusion criteria were: (1) progression to kidney failure, defined as an eGFR of less than 15 mL/ min/1.73 m<sup>2</sup>, starting on chronic dialysis, or having a kidney transplant. For the individuals with kidney failure, age at kidney failure was required, and for those who had a kidney transplant, either age at kidney failure or age at first transplant was required; (2) unrelated to all other participants up to and including the level of 3rd degree (calculated using genotype data, see Supplementary Materials); and (3) availability of single nucleotide polymorphism (SNP) array genotyping data.

#### Patient cohort descriptions

We collected 10,586 individuals from 12 cohorts from the following countries: Canada—Canadian study of prediction of death, dialysis and interim cardiovascular events (CanPREDDICT), Genetic Epidemiology of CKD Outcomes (GECKO), GECKO—Transplant (GEKCO-T); USA—Deterioration of Kidney Allograft Function (DeKAF), Genomics of kidney transplantation (GEN03); Finland—Finnish Red Cross Blood Service (FRCBS); Netherlands—TransplantLines (TL); France—Kidney Transplantation—Genomic Investigation of Essential Clinical concerns (KiT-GENIE); Ireland—Irish kidney gene project (IKGP), United Kingdom and Ireland Renal Transplant consortium (UKIRTC); UK—Salford Kidney Study (SKS), Queen's University Belfast (QUB). See Supplementary Materials for more detailed information on the recruitment and characteristics of each of these cohorts involved.

Patients were classified into primary kidney disease groups of glomerulonephritis (including IgA nephropathy), ADPKD, type 2 diabetes, and other/unknown. These groups were chosen as they represented the largest groups of patients, and other groups had too few patients (< 10%of cohort). Other (39%) and unknown (11%) forms of primary kidney disease were grouped together due to the heterogeneous nature of the other forms of primary kidney disease. This group included: hereditary nephropathy, Alport syndrome, congenital renal hypoplasia, interstitial nephritis, congenital renal dysplasia, renal vascular disease, nephrocalcinosis, and kidney tumor. If a patient was a kidney transplant recipient, and only had information on age at kidney transplant (not age at kidney failure directly), then age at transplant was used to calculate estimated age at onset of kidney failure. For such individuals, we used age at transplant minus four years for the North American cohorts and age at transplant minus 2.5 years for the European cohorts as a proxy for age at onset of kidney failure. This was based on the average wait time in the US being 3–5 years [20] and the average time in Europe being 2–3 years [21, 22].

#### Polygenic risk score calculation

Genotype data were subject to quality control for minor allele frequency, missingness per marker, and missingness per individual (see Supplementary Materials). We calculated polygenic risk score in each individual for hypertension [11], eGFR [23], rapid decline in eGFR (defined by > 5 ml/ min/1.73 m<sup>2</sup>/year) [14], albuminuria [12], and total kidney volume [13] using published data from genome-wide association studies for each trait. Further details of these genome-wide association studies can be found in Supplementary Table S2. Polygenic risk scores were calculated using PRSice2 [24], selecting alleles with a p-value threshold greater than 0.5 (see Supplementary Materials for further details). All analyses were conducted in R, using version 4.2.1 (2022-06-23) [25].

For two of these polygenic risk scores (eGFR and total kidney volume), we hypothesized that higher values would be associated with better kidney function [26], while for

 Table 1 Demographic characteristics of study participants

Cohort	Number of patients	Age at kidney failure (years), median (range)	Female, <i>n</i> (%)	Primary renal disease			
				Glomerulone- phritis, n (%)	Other/ unknown, <i>n</i> (%)	Autosomal dominant poly- cystic kidney disease, n (%)	Type 2 diabetes, n (%)
CanPREDDICT	614	70 (23–94)	217 (35)	83 (14)	247 (40)	47 (8)	237 (39)
DeKAF	1676	48 (0–79)	620 (37)	390 (23)	532 (32)	263 (16)	491 (29)
FRCBS	954	56 (18-77)	297 (31)	259 (27)	445 (47)	170 (18)	80 (8)
GECKO	84	68 (20–93)	39 (46)	7 (8)	45 (54)	6 (17)	26 (31)
GECKO-T	276	50 (1-79)	111 (10)	49 (18)	122 (44)	41 (15)	64 (23)
GEN03	998	47 (1–77)	397 (40)	293 (29)	347 (35)	139 (14)	219 (22)
KiT-GENIE	1831	52 (16-85)	644 (35)	0 (0)	1551 (85)	280 (15)	0 (0)
IKGP	194	50 (25-84)	91 (47)	0 (0)	0 (0)	194 (100)	0 (0)
QUB	150	41 (2–70)	62 (41)	32 (21)	96 (64)	18 (12)	4 (3)
SKS	732	68 (24–95)	246 (34)	61 (8)	415 (57)	89 (12)	167 (23)
TL	1056	48 (16–72)	448 (42)	264 (25)	642 (61)	129 (12)	21 (2)
UKIRTC	2021	44 (16–77)	738 (37)	318 (16)	1449 (72)	234 (12)	20(1)
Total	10,586	50 (0-95)	3910 (37)	1756 (17)	5891 (56)	1610 (15)	1329 (13)

For further details regarding the recruitment and characteristics of each cohort, see the Supplementary Materials

*CanPREDDICT* Canadian study of prediction of death, dialysis and interim cardiovascular events, *DeKAF* deterioration of kidney allograft function, *FRCBS* finnish red cross blood service, *GECKO* genetic epidemiology of CKD outcomes, *GEKCO-T* GECKO-transplant, *GEN03* genomics of kidney transplantation, *KiT-GENIE* kidney transplantation-genomic investigation of essential clinical concerns, *IKGP* Irish kidney gene project, *QUB* Queen's University Belfast, *SKS* Salford kidney study, *TL* transplantlines, *UKIRTC* United Kingdom and Ireland renal transplant consortium

the others (hypertension, albuminuria, stroke, intracranial aneurysm, and rapid kidney function decline), one might expect that higher values would be associated with worse kidney function. To simplify interpretation, we standardized the directionality of all polygenic risk scores, such that one might expect higher scores to be associated with negative outcomes. We did this by inverting the sign of the eGFR and total kidney volume polygenic risk scores to create "new" polygenic risk scores, which we will refer to as "decreased eGFR", and "decreased total kidney volume".

# **Statistical analysis**

Patients were split into two groups, "high" polygenic risk score, and "non-high" polygenic risk score, for each of the five polygenic traits for the following thresholds: high polygenic risk score—top 2%, non-high polygenic risk score—top 5%, non-high polygenic risk score—top 5%, non-high polygenic risk score—top 10%, non-high polygenic risk score—top 20%, non-high polygenic risk score at kidney failure was compared between the "high" polygenic risk score and the "non-high" polygenic risk score groups for each of these thresholds using the function t test.

Data were split into four subgroups for each primary kidney disease (glomerulonephritis, ADPKD, type 2 diabetes and other/unknown). Mean age at kidney failure for each primary kidney disease subgroup was compared between high and non-high polygenic risk score burden (defined as top 10% and bottom 90%) for each trait using a series of ttests in R.

Univariable logistic regression models were created using the R function *glm* to test the association between sex, primary kidney disease, and each polygenic risk score trait and the odds of developing kidney failure before the age of 60. The effect of each polygenic risk score was investigated as a continuous variable, and as a discrete variable (discretized into non-high and high burden in the above manner).

Multivariable logistic regression models were then created using the R function *glm* to test the association between the statistically significant factors from the univariate analysis and the odds of developing kidney failure before the age of 60. Separate models were constructed for each continuous and discrete polygenic risk score.

We also conducted power calculations to determine the smallest effect size that could be reliably detected (see Supplementary Materials). Table 2Comparison of age atkidney failure using differentdefinitions of high and non-highPRS burden for each trait

PRS	Threshold	Age at kidney failure (years)		
		High burden	Non-high burden	P value
Albuminuria	Top 2% vs bottom 98%	49.5	49.7	0.78
	Top 5% vs bottom 95%	50.1	49.7	0.59
	Top 10% vs bottom 90%	49.5	49.8	0.66
	Top 20% vs bottom 80%	49.3	49.8	0.18
Hypertension	Top 2% vs bottom 98%	48.8	49.8	0.39
	Top 5% vs bottom 95%	48.8	49.8	0.15
	Top 10% vs bottom 90%	49	49.8	0.14
	Top 20% vs bottom 80%	49.5	49.8	0.46
Decreased eGFR	Top 2% vs bottom 98%	49	49.7	0.48
	Top 5% vs bottom 95%	47.9	49.8	0.004
	<b>Top 10% vs bottom 90%</b>	47.9	49.9	0.00005
	<b>Top 20% vs bottom 80%</b>	48.8	50	0.002
Decreased total kidney volume	Top 2% vs bottom 98%	50.3	49.7	0.60
	Top 5% vs bottom 95%	50.3	49.7	0.38
	Top 10% vs bottom 90%	50.1	49.7	0.49
	Top 20% vs bottom 80%	50.2	49.6	0.16
Rapid eGFR decline	Top 2% vs bottom 98%	50.4	50.6	0.88
	Top 5% vs bottom 95%	49	50.7	0.04
	Top 10% vs bottom 90%	49.6	50.8	0.07
	Top 20% vs bottom 80%	50.2	50.7	0.28

P values < 0.05 are bolded

There is a difference of between 1.3 and 2 years in age at kidney failure between patients with high and non-high burden for decreased eGFR. Given that we tested 5 traits, a Bonferroni-adjusted significance threshold would be P < 0.01

PRS polygenic risk score, eGFR estimated glomerular filtration rate

# Results

Table 1 shows the characteristics of the 10,586 patients with kidney failure that passed genotyping Quality Control. The median age at kidney failure was 50 years and there were more males (6676, 63%) than females. One thousand six hundred and ten individuals (15%) had ADPKD, 1756 (17%) had glomerulonephritis, 1329 (13%) had type 2 diabetes, 5891 (56%) had some other or unknown form of kidney disease. Supplementary Table S1 contains patient characteristics split by primary kidney disease. Power calculations (for 95% power and 0.05 significance) showed that the smallest effect size that could be reliably detected is 8.5 months.

### Influence of polygenic burden on age at kidney failure

To investigate whether polygenic burden influences age at kidney failure, we compared age at kidney failure between those with high polygenic risk score burden for each trait for five thresholds (top 2% vs rest, top 5% vs rest, top 10% vs rest, top 20% vs rest). Out of the five studied polygenic risk score traits, only the decreased eGFR polygenic risk score showed a significant difference between high and non-high polygenic burden (Table 2). Patients in the top 5% and top 20% of decreased eGFR polygenic risk score burden developed kidney failure at a significantly younger age compared to those in the bottom 95% and 80% of burden, respectively (47.9 years versus 49.8 years; P = 0.004, 48.8 years vs 50.0 years; P = 0.002, respectively, Table 2). Similarly, patients within the top 10% of decreased eGFR polygenic risk score burden developed kidney failure at a significantly younger age compared to the bottom 90% of burden (47.9 years versus 49.9 years; P < 0.0005, Table 2, Supplementary figure S1).

We then examined the impact of the high and non-high polygenic burden of the five polygenic risk score traits on the patient's primary kidney disease. Statistically significant differences were observed in the age at kidney failure and decreased eGFR polygenic risk score burden between glomerulonephritis and type 2 diabetes (Supplementary table S3). Among patients with glomerulonephritis, those in the top 10% of decreased eGFR polygenic risk score

Journal of Nephrology (2025) 38:969–978

**Table 3** Univariable logistic regression models investigating the association between kidney failure development before the age of 60 yearsand sex, primary renal disease, continuous PRS, and discretized PRS

	OR (95% CI)	P value
Sex		
Female	1	_
Male	0.87 (0.79, 0.95)	0.001
Primary renal disease		
Glomerulonephritis	1	_
Other/Unknown	0.59 (0.51, 0.67)	2e-15
ADPKD	0.74 (0.62, 0.87)	2.9e-4
Type 2 Diabetes	0.32 (0.27, 0.38)	<2e-16
Continuous PRS		
Hypertension PRS	0.99 (0.95, 1.03)	0.75
Decreased eGFR PRS	1.05 (1.004, 1.09)	0.03
Albuminuria PRS	1.01 (0.97, 1.06)	0.6
Rapid eGFR decline PRS	1.02 (0.97, 1.07)	0.51
Decreased total kidney volume PRS	0.97 (0.93, 1.02)	0.22
Discrete PRS		
Hypertension		
Non-high (bottom 90%)	1	_
High (top 10%)	1.03 (0.89, 1.19)	0.71
Decreased eGFR		
Non-high (bottom 90%)	1	-
High (top 10%)	1.23 (1.06, 1.43)	0.006
Albuminuria		
Non-high (bottom 90%)	1	-
High (top 10%)	1.15 (0.97, 1.33)	0.06
Rapid eGFR decline		
Non-high (bottom 90%)	1	-
High (top 10%)	1.22 (1.03, 1.47)	0.03
Decreased total kidney volume		
Non-high (bottom 90%)	1	-
High (top 10%)	1.01 (0.88, 1.17)	0.88

P values < 0.05 are bolded

Given that we tested five traits, a Bonferroni-adjusted significance threshold would be P < 0.01. The odds ratio is the odds of developing kidney failure before the age of 60 years for each factor

*OR* odds ratio, *CI* confidence interval, *PRS* polygenic risk score, *eGFR* estimated glomerular filtration rate

developed kidney failure at the age of 41.6 years compared to 45.8 years for those in the bottom 90% of polygenic risk score (P=0.0007). Similarly, for patients with type 2 diabetes as their primary kidney disease, those in the top 10% of decreased eGFR polygenic risk score developed kidney failure at 52.2 years compared to 56.4 years for those in the bottom 90% of polygenic risk score (P=0.002).

To investigate which factors impact age at kidney failure, we created univariable logistic regression models for kidney failure before the age of 60 years. Male sex (Odds ratio (OR)=0.87; 95% Confidence Interval (CI) 
 Table 4
 Multivariable logistic regression models investigating the association between kidney failure development before the age of 60 years and sex, primary renal disease, continuous decreased eGFR PRS, and discretized decreased eGFR PRS

Characteristic	OR (95% CI)	P value
Model 1 (Continuous PRS)		
Sex		
Female	1	-
Male	0.87 (0.79-0.95)	0.002
Primary renal disease		
Glomerulonephritis	1	-
Type 2 diabetes	0.32 (0.27-0.38)	<2e-16
ADPKD	0.72 (0.61-0.85)	1.3e-4
Other/Unknown	0.58 (0.50-0.66)	4e-16
Decreased eGFR PRS	1.05 (1.01–1.10)	0.01
Model 2 (Discrete PRS)		
Sex		
Female	1	-
Male	0.87 (0.79-0.95)	0.002
Primary renal disease		
Glomerulonephritis	1	-
Type 2 diabetes	0.32 (0.27-0.38)	<2e-16
ADPKD	0.72 (0.61-0.85)	1.4e-4
Other/Unknown	0.58 (0.51-0.66)	5e-16
Decreased eGFR PRS		
Non-high (bottom 90%)	1	-
High (top 10%)	1.26 (1.08–1.46)	0.003

P values < 0.05 are bolded

Given that we tested 5 traits, a Bonferroni-adjusted significance threshold would be P < 0.01

The odds ratio is the odds of developing kidney failure before the age of 60 years for each factor

*OR* odds ratio, *CI* confidence interval, *PRS* polygenic risk score, *eGFR* estimated glomerular filtration rate

0.79–0.95; P = 0.001, Table 3) and primary kidney disease type were both significantly associated with kidney failure development. Decreased eGFR polygenic risk score as a continuous variable was associated with kidney failure before the age of 60 years (OR = 1.05; 95% CI 1.004–1.09; P = 0.03, Table 3), but did not withstand a multiple testing correction. The discretized eGFR polygenic risk score has a stronger and more significant association with kidney failure before the age of 60 years (OR = 1.23; 95% CI 1.06–1.43; P = 0.006, Table 3). Rapid eGFR decline polygenic risk score did not pass the multiple correction threshold of 0.01 (OR = 1.22; 95% CI 1.03–1.47; P = 0.03, Table 3). There was no significant effect of polygenic risk score for hypertension, albuminuria, or decreased total kidney volume burden.

Two separate multivariable logistic regression models were then created to predict kidney failure development before the age of 60 years, both controlling for sex and primary kidney disease, one with the continuous decreased eGFR polygenic risk score, and the other with the discrete one. Male sex and primary kidney disease were both significant in these models, as was the continuous decreased eGFR polygenic risk score (OR = 1.05; 95% CI 1.01–1.10; P = 0.01, Table 4), and the discrete decreased eGFR polygenic risk score (OR = 1.26; 95% CI 1.08–1.46; P = 0.003, Table 4).

### Discussion

We explored the impact of polygenic burden for kidney traits on age at kidney failure using 12 cohorts comprising 10,586 individuals with kidney failure.

A top 10th percentile polygenic risk score for decreased eGFR was associated with two years earlier age at kidney failure, and increased likelihood of developing kidney failure before the age of 60 years. This effect is seen particularly strongly in those with primary kidney diseases of glomerulonephritis and type 2 diabetes. Importantly, although hypertension and proteinuria are known to be significant risk factors for progression of CKD, we were unable to demonstrate an effect of polygenic risk score for these variables in this large cohort of CKD patients.

A previous study, involving nearly 9000 individuals, assessed the impact of polygenic risk score for eGFR on the risk of developing both CKD and kidney failure and found hazard ratios of 1.30 and 1.20, respectively associated with standard deviation increases in polygenic risk scores [19]. However, only 470 individuals (5%) had kidney failure, significantly less than our over 10,000 individuals with kidney failure. While the impact of polygenic risk score on age at kidney failure in our study is relatively modest (2 years between high and non-high polygenic burden groups), these results align with this previous study. In a separate, cross ancestry study, those in the top 2% of a CKD polygenic risk score were shown to have a nearly threefold increased risk of CKD [16]. This study utilized a different polygenic risk score which was a derivative of the eGFR polygenic risk score but was optimized for studying all global ancestry populations, not just those of European ancestry.

Furthermore, polygenic risk scores provide value in interpreting family history data [27]. Polygenic risk scores have been developed for different kidney disease types and, in particular, they are most advanced for IgA nephropathy where a recent genome-wide association study of more than 10,000 individuals from 17 international cohorts identified 30 loci associated with IgA nephropathy development, and these loci explained 11% of disease risk [28]. In contrast to monogenic disease caused by rare pathogenic variants, where genetic and allelic heterogeneity is the expectation, polygenic scores seek to identify the cumulative impact of common variants of small individual effects which would be expected to contribute to disease progression regardless of the primary etiology. These variants may work through a multitude of pathophysiological processes including, but not limited to, fibrosis, inflammation, metabolic disease, hypertension, and more. Whilst polygenic risk scores typically explain only a fraction of the heritability of a trait, this is likely to improve as genome-wide association study sample sizes increase and methods are developed to integrate rare variants into polygenic risk scores [29]. It is also important to note that the polygenic risk score for rapid eGFR decline was calculated using genome-wide association studies involving 19,000 cases and 175,000 controls, that only found 7 genome-wide significant loci (variance explained not reported), compared to over 260 loci for the eGFR genome-wide association studies, involving over 560,000 individuals and explaining 7.1% of trait variation (Supplementary Table S2). This relative lack of power of the underlying genome-wide association studies may partially explain why we did not find significant associations of age at kidney failure with this rapid eGFR decline polygenic risk score.

A recent study with more than 12,000 multiple sclerosis patients found that the variants associated with disease severity are different from those associated with disease susceptibility [30]. Such a study highlights the need for specific, well-powered genome-wide association studies for kidney failure, thus enabling more powerful kidney failure polygenic risk scores. The genome-wide association study for eGFR, upon which this study is based, was conducted in largely healthy individuals. It is possible that significantly different variants may play a role in a genome-wide association study of kidney failure than a genome-wide association study for eGFR in largely healthy individuals.

Although the effect sizes observed in our study may be considered modest, they are statistically significant and can explain two years difference in age at kidney failure. Furthermore, there was a 26% greater risk of kidney failure before the age of 60 years in the high polygenic risk score group (top 10%) compared to the bottom 90% polygenic risk score group. As genome-wide association studies continue to grow in size and predictive power, polygenic risk score could potentially explain a more substantial proportion of the age at kidney failure, thereby becoming increasingly valuable in clinical decision-making. In such a case, polygenic risk score for decreased eGFR could be one of a number of factors used to aid in risk stratification and personalized treatment decisions (including proactive disease management) for individuals at risk of kidney failure.

Our study has limitations. Age at kidney failure was not known in all individuals and we used proxy ages at kidney failure in some of the cohorts with transplant recipients. Broad categorization of primary kidney diseases also introduced noise and reduced statistical power (the smallest effect size that this study is powered to detect is approximately 8.5 months). Future studies should strive for more precise phenotyping. It is notable that in this cohort, male sex results in a 13% lower risk of developing kidney failure before the age of 60 compared to females. This result is not actually necessarily contradictory to the fact that men are at higher risk of kidney failure than women. We found this in our cohort, in that 65% of the kidney failure cohort is male. Additionally, while our dataset comprises a diverse range of kidney diseases, it is possible that polygenic risk score performance may vary across different kidney disease types. While our cohort provides some evidence of this possibility, more research is required. Other clinical factors that may modify the effect, such as hypertension and diabetes status, were not available and may have improved prediction.

It is apparent from our study that the utility of polygenic risk score in defining renal phenotype appears to be primarily at the extremes of polygenic risk score, i.e., when comparing patients with high polygenic risk score for each variable compared to those with non-high polygenic risk score. Similar effects have also been found by others including one study which reported individuals in the top 2% of CKD polygenic risk score as having a nearly threefold increased risk of CKD compared to those in the bottom 98% [16]. Another large study reported that polygenic risk alters the penetrance of monogenic kidney disease [18]. It is notable that, while our results suggest those in the top 2% have an earlier age at kidney failure than those in the bottom 98% (49 years vs 49.7 years), this is not statistically significant (P = 0.48), likely due to a lack of power involved by comparing a group with relatively few individuals.

Given that our dataset exclusively consisted of individuals of European ancestry, caution should be exercised when applying our findings to other populations. Studies like those that create a trans-ancestry polygenic risk score [16] may offer valuable insights into diverse patient groups.

In summary, our meta-analysis unveils the impact of inherited genetic factors associated with kidney function on age at kidney failure. The incorporation of the kidney function-related polygenic risk score into clinical practice may hold promise for risk assessment and treatment strategies. While challenges remain, the growing power of gene wide association studies and polygenic risk score warrants further investigation into their potential utility in improving patient outcomes and informing transplantation decisions. This is made increasingly feasible by the ever-decreasing cost of single nucleotide polymorphism-array genotyping, which now costs in the region of \$50 per sample [31]. These results suggest that genetic factors related to kidney function play a role in determining the timing of kidney failure. Our findings provide novel insights into the genetic determinants of kidney failure onset and have potential implications for clinical practice and future research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40620-025-02207-7.

Author contributions The authors confirm contribution to the paper as follows: study conception and design: KEC, EG, MBL, PJC, GLC; data collection: KEC, EG, VM, PG, KAB, EAEE, SM, CH, AJMcK, APM, PJM, MHB, WG, PAJ, AKI, BJK, GML, SM, IH, KH, JP, HY, PCH, JS, SS, PAK, MBL, SL; analysis and interpretation of results: KEC, EG, MBL, GLC, PJC, SM; draft manuscript preparation: KC, EG, GLC, PJC. All authors reviewed the results and approved the final version of the manuscript.

Funding Open Access funding provided by the IReL Consortium. This study makes use of transplant data from a number of cohorts, all described in the Methods. We thank all of our collaborators for access to their data. This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) under Grant number 18/CRT/6214, and funds from the Royal College of Surgeons in Ireland StAR PhD. CanPREDDICT, GECKO-T, and GECKO were funded by the Canadian Institutes of Health Research Project Grant (application # 427810) to MBL. The DeKAF and GEN03 cohorts were generated with support from NIH/NIAID grants 5U19-AI070119, K01AI130409, and 5U01-AI058013. FRCBS cohort study has been funded by the Government of Finland VTR funding, Munuaissäätiö (to SM), Suomen Transplantaatiokirurginen yhdistys ry (to SM), and Academy of Finland (to JP). The genotype imputation reference panel used for the research were obtained from THL Biobank (study number: BB2019\_12). We thank all study participants for their generous participation in biobank research. The KiT-GENIE cohort has been funded by several entities: (1) The Etoiles Montantes funding by the Pays de la Loire region (n°2018-09998), (2) The IRCT Dialyse research project by the Société Francophone de Néphrologie, Dialyse et Transplantation (SFNDT), (3) The Greffe research project by French Agence de la Biomédecine (ABM, n°18GREFFE014). In addition, we thank the Roche Pharma, Novartis, Astellas, Chiesi, Sandoz and Sanofi laboratories for supporting the DIVAT (KiT-GENIE) cohort as the CENTAURE Foundation (http://www.fondation-centaure.org). The IKGP is funded by the Irish Kidney Foundation and was supported by a grant from the HRCI-HRB Joint funding scheme (Grant code: HRCI-HRB-2020-032). The QUB contribution was supported by the HSC R&D division (STL/5569/19), UKRI (MRC MC\_PC\_20026), and Science Foundation Ireland-NI Department for the Economy Investigator Program Partnership Award (15/IA/3152). QUB data was generated using NICOLA (Northern Ireland Cohort for the Longitudinal Study of Ageing) data; NICOLA's molecular data was supported by The Atlantic Philanthropies and the Economic and Social Research Council (ES/L008459/1). SKS reports funding support from Shire, CSL Vifor, BergenBio and Astellas. TL was supported by a grant from Astellas BV. The UKIRTC was supported by grants awarded from the Wellcome Trust (090355/A/09/Z, 090355/B/09/Z and 088849/Z/09/Z, "WTCCC3"), the Medical Research Council (grants G0600698 and MR/J006742/1 to S.H.S.; G0802068 to G.M.L. and MR/K002996/1 to G.M.L; grants G0801537/ID: 88245), Guy's and St Thomas' Charity (grants R080530 and R090782) to M.H.F. and G.M.L., from the European Union FP7 (grant agreement no HEALTHF5-2010-260687 to M.H.F. and project number 305147: BIO-DrIM to M.H.F. and I.R.M.); and by the National Institute for Health Research Biomedical Research Centre at Guy's and St Thomas' and King's College London.

**Data availability** The data used in this study are not publicly available due to concerns regarding patient confidentiality and privacy.

Access to the data can be requested from each participating center individually, subject to their respective data sharing policies and ethical considerations.

### Declarations

**Conflict of interest** Unless otherwise stated all authors have no conflicts of interest. IH reports receiving research funding from MSD and Hansa Biopharma, and has ongoing consultancy agreements with Novartis, Hansa Biopharma, and Takeda. MBL has received advisory and consulting fees from Bayer, Sanofi, Reata, GlaxoSmithKline, and Otsuka. GLC and PJC report receiving research funding relating to kidney disease genetics from Biomarin.

**Ethical approval** This study, which involved human participants, their data or biological material, was approved (or granted exemption) by the appropriate institutional and/or national research ethics committee which was the Hammersmith and Queen Charlotte's & Chelsea Research Ethics Committee REC No 08/H0707/1.

**Informed consent to participate** All participants have provided their informed consent to participate and publish this research.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

# References

- Kovesdy CP (2022) Epidemiology of chronic kidney disease: an update 2022. Kidney Int Suppl 12(1):7–11. https://doi.org/10. 1016/j.kisu.2021.11.003
- United States Renal Data System. 2023 USRDS Annual Data Report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD. Published online 2023.
- 3. Vaidya S, Aeddula N (2022) Chronic kidney disease. StatPearls Publishing
- Staples A, Wong C (2010) Risk factors for progression of chronic kidney disease. Curr Opin Pediatr 22(2):161–169. https://doi.org/10.1097/MOP.0b013e328336ebb0
- Xie X, Liu Y, Perkovic V et al (2016) Renin-angiotensin system inhibitors and kidney and cardiovascular outcomes in patients with CKD: A Bayesian network meta-analysis of randomized clinical trials. Am J Kidney Dis 67(5):728–741. https://doi.org/ 10.1053/j.ajkd.2015.10.011
- Yau K, Dharia A, Alrowiyti I, Cherney DZI (2022) Prescribing SGLT2 inhibitors in patients with CKD: expanding indications and practical considerations. Kidney Int Rep 7(7):1463–1476. https://doi.org/10.1016/j.ekir.2022.04.094
- Connaughton DM, Bao MBB, Kennedy C et al (2019) Monogenic causes of chronic kidney disease in adults. Kidney Int 95(4):914–928. https://doi.org/10.1016/j.kint.2018.10.031

- Jefferis J, Hudson R, Lacaze P et al (2023) Monogenic and polygenic concepts in chronic kidney disease (CKD). J Nephrol. https://doi.org/10.1007/s40620-023-01804-8
- Satko SG, Freedman BI (2005) The familial clustering of renal disease and related phenotypes. Med Clin North Am 89(3):447– 456. https://doi.org/10.1016/j.mcna.2004.11.011
- Tin A, Köttgen A (2020) Genome-wide association studies of CKD and related traits. Clin J Am Soc Nephrol 15(11):1643– 1656. https://doi.org/10.2215/CJN.00020120
- Bi W, Fritsche LG, Mukherjee B, Kim S, Lee S (2020) A fast and accurate method for genome-wide time-to-event data analysis and its application to UK Biobank. J Clean Prod 107(2):222– 233. https://doi.org/10.1016/j.ajhg.2020.06.003
- Teumer A, Li Y, Ghasemi S et al (2019) Genome-wide association meta-analyses and fine-mapping elucidate pathways influencing albuminuria. Nat Commun. https://doi.org/10.1038/ s41467-019-11576-0
- Liu Y, Basty N, Whitcher B et al (2021) Genetic architecture of 11 organ traits derived from abdominal MRI using deep learning. Elife 10:1–30. https://doi.org/10.7554/eLife.65554
- Gorski M, Jung B, Li Y et al (2021) Meta-analysis uncovers genome-wide significant variants for rapid kidney function decline. Kidney Int 99(4):926–939. https://doi.org/10.1016/j. kint.2020.09.030
- Patel AP, Wang M, Ruan Y et al (2023) A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. Nat Med 29(7):1793–1803. https://doi.org/10.1038/ s41591-023-02429-x
- Khan A, Turchin MC, Patki A et al (2022) Genome-wide polygenic score to predict chronic kidney disease across ancestries. Nat Med 28(7):1412–1420. https://doi.org/10.1038/ s41591-022-01869-1
- Bakshi A, Jefferis J, Wolfe R et al (2023) Association of polygenic scores with chronic kidney disease phenotypes in a longitudinal study of older adults. Kidney Int 103(6):1156–1166. https://doi.org/10.1016/j.kint.2023.03.017
- Khan A, Shang N, Nestor JG et al (2023) Polygenic risk alters the penetrance of monogenic kidney disease. Nat Commun 14(1):1–10. https://doi.org/10.1038/s41467-023-43878-9
- Yu Z, Jin J, Tin A et al (2021) Polygenic risk scores for kidney function and their associations with circulating proteome, and incident kidney diseases. JASN 32(12):3161–3173. https://doi. org/10.1681/ASN.2020111599
- 20. Understanding the transplant waitlist. National Kidney Foundation. https://www.kidney.org/content/understanding-transplantwaitlist
- 21. How long is the wait for a kidney? Organ transplantation - NHS Blood and Transplant. https://www.nhsbt. nhs.uk/organ-transplantation/kidney/receiving-a-kidney/ how-long-is-the-wait-for-a-kidney/
- Smyth G, O'Seaghdha C, O'Kelly P, Cooney A (2023) National Kidney Transplant Service Annual Report 2023. National Kidney Transplant Service. http://www.beaumont.ie/media/NationalKi dneyTransplantServiceAnnualReport20231.pdf
- Wuttke M, Li Y, Li M et al (2019) A catalog of genetic loci associated with kidney function from analyses of a million individuals. Nat Genet 51(6):957–972. https://doi.org/10.1038/ s41588-019-0407-x
- Choi SW, Mak TSH, O'Reilly PF (2020) A guide to performing polygenic risk score analyses. Nat Protoc 15(9):2759–2772. https://doi.org/10.1038/s41596-020-0353-1
- 25. Computing RF for S. R: A language and environment for statistical computing (2021) Published online
- Wu J, Wang Y, Vlasschaert C et al (2024) Kidney volume and risk of incident kidney outcomes. JASN. https://doi.org/10.1681/ASN. 000000000000419

- Mars N, Lindbohm JV, della Briotta Parolo P, et al (2022) Systematic comparison of family history and polygenic risk across 24 common diseases. Am J Hum Genet 109(12):2152–2162. https://doi.org/10.1016/j.ajhg.2022.10.009
- Kiryluk K, Sanchez-Rodriguez E, Zhou XJ et al (2023) Genomewide association analyses define pathogenic signaling pathways and prioritize drug targets for IgA nephropathy. Nat Genet 55(7):1091–1105. https://doi.org/10.1038/s41588-023-01422-x
- Aldisi R, Hassanin E, Sivalingam S et al (2023) Gene-based burden scores identify rare variant associations for 28blood biomarkers. BMC Genom Data 24:50. https://doi.org/10.1186/ s12863-023-01155-0
- 30. Harroud A, Stridh P, International Multiple Sclerosis Genetics Consortium et al (2023) Locus for severity implicates

CNS resilience in progression of multiple sclerosis. Nature 619(7969):323–331. https://doi.org/10.1038/s41586-023-06250-x

 Pavan S, Delvento C, Ricciardi L, Lotti C, Ciani E, D'Agostino N (2020) Recommendations for choosing the genotyping method and best practices for quality control in crop genome-wide association studies. Front Genet 11:447. https://doi.org/10.3389/fgene. 2020.00447

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# **Authors and Affiliations**

Kane E. Collins<sup>1,2,3</sup> · Edmund Gilbert<sup>1,2</sup> · Vincent Mauduit<sup>4</sup> · Pukhraj Gaheer<sup>5</sup> · Elhussein A. E. Elhassan<sup>6,7</sup> · Katherine A. Benson<sup>1,2</sup> · Shohdan Mohamad Osman<sup>6,7</sup> · Claire Hill<sup>8</sup> · Amy Jayne McKnight<sup>8</sup> · Alexander Peter Maxwell<sup>8</sup> · Peter J. van der Most<sup>9</sup> · Martin H. de Borst<sup>10</sup> · Weihua Guan<sup>11</sup> · Pamala A. Jacobson<sup>12</sup> · Ajay K. Israni<sup>13</sup> · Brendan J. Keating<sup>14</sup> · Graham M. Lord<sup>15</sup> · Salla Markkinen<sup>16</sup> · Ilkka Helanterä<sup>17</sup> · Kati Hyvärinen<sup>16</sup> · Jukka Partanen<sup>16</sup> · Stephen F. Madden<sup>18</sup> · Joshua Storrar<sup>19,20</sup> · Smeeta Sinha<sup>19,20</sup> · Philip A. Kalra<sup>19,20</sup> · Matthew B. Lanktree<sup>5</sup> · Sophie Limou<sup>4</sup> · Gianpiero L. Cavalleri<sup>1,2,3</sup> · Peter J. Conlon<sup>5,6</sup>

- Peter J. Conlon peterconlon@beaumont.ie
- <sup>1</sup> School of Pharmacy and Biomolecular Sciences, Royal College of Surgeons in Ireland, Dublin, Ireland
- <sup>2</sup> The Science Foundation Ireland FutureNeuro Centre of Excellence, Dublin, Ireland
- <sup>3</sup> SFI Centre for Research Training in Genomics Data Science, University of Galway, Galway, Ireland
- <sup>4</sup> Nantes University, Ecole Centrale Nantes, INSERM, Center for Research in Transplantation and Translational Immunology, UMR1064, Nantes, France
- <sup>5</sup> Division of Nephrology, Departments of Medicine and Health Research Methodology, Evidence and Impact, St. Joseph's Healthcare Hamilton, McMaster University and Population Health Research Institute, Hamilton, ON, Canada
- <sup>6</sup> Department of Nephrology and Transplantation, Beaumont Hospital, Dublin, Ireland
- <sup>7</sup> Department of Medicine, Royal College of Surgeons in Ireland, Dublin, Ireland
- <sup>8</sup> Centre for Public Health, Queen's University Belfast, Belfast, UK
- <sup>9</sup> Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

- <sup>10</sup> Department of Internal Medicine, Division of Nephrology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
- <sup>11</sup> Division of Biostatistics and Health Data Science, School of Public Health, University of Minnesota, Minneapolis, MN, USA
- Experimental and Clinical Pharmacology, College of Pharmacy, University of Minnesota, Minneapolis, MN, USA
- <sup>13</sup> University of Texas Medical Branch, Galveston, TX, USA
- <sup>14</sup> Department of Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
- <sup>15</sup> School of Immunology and Microbial Sciences, University College London, London, UK
- <sup>16</sup> Finnish Red Cross Blood Service, Research and Development, Biomedicum 1, Helsinki, Finland
- <sup>17</sup> Helsinki University Hospital, Transplantation and Liver Surgery, Helsinki, Finland
- <sup>18</sup> Data Science Centre, Beaux Lane House, Royal College of Surgeons in Ireland, Dublin, Ireland
- <sup>19</sup> Salford Royal Hospital, Northern Care Alliance NHS Foundation Trust, Salford, UK
- <sup>20</sup> University of Manchester, Manchester, UK