
Article:
Chen, E.W.C., Chong, J., Valluru, M.K. orcid.org/0000-0001-9156-866X et al. (4 more authors) (2024) Combining genotype with height-adjusted kidney length predicts rapid progression of ADPKD. Nephrology Dialysis Transplantation, 39 (6). pp. 956-966. ISSN 0931-0509

https://doi.org/10.1093/ndt/gfad270

The combination of genotype and height adjusted kidney length improves risk prediction of rapid disease progression in autosomal dominant polycystic kidney disease

Eugene WC Chen[1,2]

Jiehan Chong[1,2]

Manoj K Valluru[1]

Miranda Durkie[3]

Roslyn J Simms[1,2]

Peter C Harris[4]

Albert CM Ong[1,2]

[1]Academic Nephrology Unit, Division of Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Beech Hill Road, Sheffield, United Kingdom

[2]Sheffield Kidney Institute, Sheffield Teaching Hospitals NHS Foundation Trust, Herries Road, Sheffield, United Kingdom

[3]Sheffield Diagnostics Genetic Service, North East and Yorkshire Genomic Laboratory Hub, Sheffield Children's NHS Foundation Trust, Sheffield, UK

[4]Division of Nephrology and Hypertension, Mayo Clinic and Foundation, Rochester, USA

Correspondence:

Prof Albert CM Ong

Division of Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Beech Hill Road, Sheffield S10 2RX, United Kingdom

a.ong@sheffield.ac.uk

Running head: Real world risk prediction in ADPKD

Summary:

Introduction: Our main objective was to identify baseline prognostic factors predictive of rapid disease progression in a large unselected clinical ADPKD cohort.

Methods: A cross-sectional analysis was performed in 618 consecutive ADPKD patients assessed and followed-up for over a decade. 123 patients (19.9%) had reached kidney failure by the study date. Data was available for the following: baseline eGFR (n= 501), genotype (n=549), baseline ultrasound mean kidney length (MKL, n=424), height adjusted baseline MKL (htMKL, n=377). Rapid disease progression was defined as an annualised eGFR decline (ΔeGFR) of >2.5ml/min/year by linear regression over 5 years (n=158). Patients were further divided into slow, rapid and very rapid ΔeGFR classes for analysis. Genotyped patients were classified into several categories: *PKD1* (T, truncating or NT, non-truncating), *PKD2*, other genes (non-*PKD1* or *PKD2*), NMD (no mutation detected) or variants of uncertain significance (VUS).

Results: A *PKD1-T* genotype had the strongest influence on the probability of reduced baseline kidney function by age. A multivariate logistic regression model identified *PKD1-T* genotype and htMKL (>9.5 cm/m) as independent predictors for rapid disease progression. The combination of both factors increased the positive predictive value (PPV) for rapid disease progression over age 40 years and of reaching kidney failure by age 60 years to 100%. Exploratory analysis in a subgroup with available total kidney volumes (TKV) showed higher PPV (100% v 80%) and NPV (42% v 33%) in predicting rapid disease progression compared to the Mayo Imaging Classification (1C-E).

Conclusion: Real-world longitudinal data confirms the importance of genotype and kidney length as independent variables determining ΔeGFR. Individuals with the highest risk of rapid disease progression can be positively selected for treatment based on this combination.

Key learning points

What is already known about this subject:

1. A *PKD1* truncating mutation is associated with the earliest age of onset of kidney failure among different genotype groups.
2. The influence of other non-allelic factors on the individual phenotype is illustrated by significant intrafamilial variability in pedigrees including those with *PKD1* truncating variants.
3. These factors are partly but not completely accounted for in two current prognostic tools based on TKV (Mayo Imaging Classification) or genotype (PROPKD score).

What this study adds:

1. The combination of HtMKL to genotype increased the sensitivity and specificity of identifying patients with rapid disease progression in a real-world ADPKD cohort.
2. There was a high prevalence of genetically unresolved patients (24%) compared with other published cohorts.
3. PKD patients with NMD had features suggestive of a good prognosis ie more benign course for disease progression and kidney failure compared to patients with known genotypes.

What impact this may have on practice or policy:

1. Using two defined baseline factors, patients at risk of rapid disease progression can be positively identified and benefit from earlier treatment.
2. A HtMKL cut-off (9.5cm/m) alone provided high PPV and NPV for developing kidney failure by age 60 years and could be used for identifying patients at risk of rapid disease progression in lower resource systems.
3. Defining these factors will facilitate the refinement of individual risk prediction in future studies.

Keywords:

- ADPKD, PKD1, kidney length, total kidney volume, progression

Introduction

Autosomal dominant polycystic kidney disease (ADPKD) is the most common monogenic cause of kidney failure and accounts for 8-10% of prevalent patients on kidney replacement therapy (RRT) [1]. The clinical point prevalence of ADPKD has been estimated at less than 1 in 2500 [2] though more recent population-based estimates of genetic prevalence indicate frequencies of ~1 in 1000 [3], similar to earlier prevalence studies which included autopsy information [4].

Variants in two genes, *PKD1* and *PKD2,* are reported to account for >90% of ADPKD in several cohorts [5]. Several other cystic genes have since been identified in ADPKD patients without *PKD1* or *PKD2* variants: these include *GANAB, DNAJB11, IFT140, ALG5, ALG8, ALG9, PKHD1* [6-8]. In almost all the newer gene variants however, kidney disease tends to be atypical in presentation and kidney failure either absent or late in onset.

The licensing and regulatory approval of tolvaptan for ADPKD patients with either evidence or risk of rapid disease progression has led to a major step-change in disease management [9, 10]. Current evidence for rapid disease progression has been defined as a ΔeGFR >2.5 ml/min/year, ideally measured over at least 5 years [11, 12]. In the absence of ΔeGFR information especially in younger patients or those presenting late, the use of prognostic tools such as the Mayo Imaging Classification (MIC) [13] or the Predicting Renal Outcome in Polycystic Kidney Disease (PROPKD) score [14] have been recommended. The former relies on measurements of total kidney volume (TKV) usually by MRI while the latter is heavily dependent on genotype. Other proposed measures include the measurement of mean kidney lengths (MKL) by ultrasound (as a surrogate for TKV) [15] and the age of onset of kidney failure of affected relatives where known (as a surrogate for genotype) [16].

TKV is arguably the most accurate prognostic variable currently available for ADPKD though it has not yet been widely adopted in routine clinical practice [17]. The main objective of this study was to assess what other baseline factors could most accurately predict ΔeGFR in an unselected ADPKD cohort assessed and followed-up for over a decade. As secondary objectives, we report the clinical presentations, genetic architecture and phenotypic variability of this real-world cohort.

Methods (Detailed methods in Supplementary material)

A retrospective cross-sectional service evaluation was performed on all ADPKD patient referrals (n=618) assessed and managed through a specialist PKD clinic between 2010 and 2021 at Sheffield Teaching Hospitals. The mean follow-up duration was 11.11 (± 11.20) years. 123 patients (19.9%) had reached kidney failure by the study date (1 July 2021).

Demographics, diagnosis and clinical assessment

The baseline characteristics of all 618 patients are summarised in **Table 1**. Clinical characteristics were collected by standard history taking and physical examination by two nephrologists (AO, RS). Laboratory and radiological findings were recorded where available.

In those with a family history, ADPKD was diagnosed on ultrasound scans according to the Pei-Ravine criteria [16] with genetic testing performed in cases of diagnostic uncertainty. In those without a family history, we used a cut-off of >10 kidney cysts [18] by any imaging modality. In cases with less than 10 kidney cysts, we included those with >10 liver cysts to capture potential variants in the newer ADPKD genes.

Genetic testing

Genetic analysis was performed in 549 patients (89%) following individual consent initially by Sanger sequencing for *PKD1* and *PKD2* (using LR-PCR), later by using custom next generation sequencing (Sheffield and Rochester) [19] and latterly by whole genome sequencing (Genomics England). DNA samples were unavailable for analysis in 11%. Patients were classified into those with *PKD1* (T, truncating or NT, non-truncating), *PKD2*, other genes (non-*PKD1 or PKD2*), NMD (no mutation detected), VUS (variant of uncertain significance) or CGI (complex gene inheritance). Patients with a VUS (predominantly *PKD1*) were classified based on current ACMG criteria [20]. Recurrent variants were defined as those present in at least two unrelated pedigrees.

Imaging

Abdominal ultrasound scans requested for routine clinical assessment were performed by trained sonographers. Recorded baseline renal lengths from abdominal ultrasounds were available for 424 patients, 345 with known genotypes (**Figure 1**). MKL was calculated by taking the average of the left and right bipolar kidney length (cm). HtMKL was derived in patients with height data recorded (n=377). TKV was measured in a subset of patients with abdominal MRI (n=35) or CT (n=9) scans.

Kidney function, rate of eGFR decline and end stage renal disease

The estimated glomerular filtration rate (eGFR) was calculated from the serum creatinine using the CKD-EPI equation [21]. Rapid disease progression was defined as a ΔeGFR >2.5 ml/min/year based on the median ΔeGFR reported in patients with Mayo Class 1C [11]. We defined a ΔeGFR <2.5 ml/min/year as 'slow' and a ΔeGFR >5 ml/min/year as 'very rapid' disease progression. Kidney failure was defined as an eGFR <15ml/min/1.73m^2 or the onset of RRT by the study date (1 July 2021).

Correlations of baseline factors with the rate of eGFR decline

Genotype information (*PKD1-T, PKD1-NT* and *PKD2*), HtMKL and ΔeGFR were available in a subset of patients (n = 70). Univariate logistic regression was performed to determine the odds ratio (OR) of genotypes (*PKD1-T, PKD1-NT* or *PKD2*), HtMKL >9.5cm/m, MKL >16.5cm, age at diagnosis ≤35 years or age at first clinic presentation ≤46 years with ΔeGFR. A HtMKL cut-off of >9.5cm/m was derived by dividing the MKL cut-off of >16.5cm by the mean height (1.73m) of our cohort [15]. The age cut-offs ≤35 years and ≤46 years were the median ages of the cohort for age of diagnosis and age at 1st clinic presentation respectively.

A multivariate logistic regression model was generated using baseline factors significant on univariate analysis. The positive predictive value (PPV), negative predictive value (NPV), sensitivity and specificity of each model was calculated. Further subgroup analyses were conducted based on baseline CKD stages (1-2 v 3-4), age at presentation (< >40years), ΔeGFR >5ml/min/year or with TKV (MIC Class 1C-E). In the subset of patients reaching kidney failure, the model was tested using an age cut-off of 60 years similar to the PROPKD score [14].

Statistical analysis

All data were analysed using the PRISM software package. Data were reported as mean ± SD for normally distributed data or median and interquartile range (IQR) for skewed data. Comparisons between groups were done by Student's t-test (two groups) or ANOVA (more than two groups) for normally distributed data and by Mann-Whitney test or the Kruskal–Wallis test as appropriate for skewed data. Kaplan–Meier analysis (log-rank) was used to compare the median age of kidney disease onset between genotypes. Categorical variables such as gender distribution were analysed using χ^2 tests. A probability value <0.05 was considered to be statistically significant.

Results

Demographics and clinical presentation

Table 1 summarises the demographic data and clinical characteristics of our cohort. The mean follow-up duration was 11.1 (±11.2) years. Mean age at diagnosis was 37.1 (± 19.6) years old with the majority of patients presenting between the ages of 20-60 years (**Figure 2A**). There was an equal gender ratio and 94% of patients were of white ethnicity. A positive family history of ADPKD was reported by 73% and was associated with a significantly younger age of diagnosis compared to those without a family history (median age of 31.0 vs. 47.0 years old; $p<0.0001$) potentially due to early screening.

37% of the patients were asymptomatic and diagnosed by cascade screening due to a positive family history. Of those symptomatic at presentation (44%), lumbar/abdominal pain was the most frequent symptom (40%) and hypertension (20%) was the most common feature. All other patients (18%) were diagnosed incidentally through abnormal imaging (92%) or deranged blood biochemistry (increased creatinine or liver function tests) (8%). Overall, the clinical features were similar to those reported from historical ADPKD cohorts [22, 23] with some exceptions.

Hypertension was the most common clinical finding (65%) in our cohort, with a high prevalence in young patients <20 years (21.3%) which increased to 100% in those >80 years (**Figure S1A**) confirming the high incidence of hypertension in young adults with ADPKD from other cohorts [24, 25] and as a prognostic indicator of disease severity [26, 27]. Patients with hypertension had more advanced disease as reflected by a lower baseline eGFR and greater HtMKL at presentation and its severity (number of antihypertensive drugs prescribed) correlated with HtMKL (**Figure S1B-D)**.

Associations with gender

Despite a similar ΔeGFR, males presented at a significantly older age (median: 47 vs 42 years; $p<0.05$) and with a lower baseline eGFR (median eGFR: 63.9ml/min/1.73m^2 vs 83.7ml/min/1.73m^2; $p<0.05$) than females. The presence of an abdominal 'mass' was a more common initial presentation in females ($p<0.05$) while the diagnosis was more commonly made incidentally in males. Hypertension and gout were more common in males, while urinary tract infections occurred more frequently in females ($p<0.05$).

Genotypes

Genotyping information was available for 89% (of the cohort as shown in **Figure 3A**. *PKD1* variants were the most common (52%), followed by *PKD2* (19%). 59% *PKD1* variants were classified as protein truncating variants (T) while 41% were predicted to be non-truncating (NT). Other mutation categories included other genes (n=20), variants of uncertain significance (VUS, n=38) or a complex genetic inheritance (mosaic, biallelic, digenic, n=7). No pathogenic variants were found in 17%. Full details of other pathogenic alleles, VUS variants and complex gene inheritance can be found in **Supplementary Tables 1-3.**

The vast majority of VUS changes were in *PKD1* (**Figure 3B**). The *PKD1* VUS and *PKD2* VUS groups were comparable in HtMKL, MKL, ΔeGFR, baseline hypertension and age at presentation to their pathogenic counterparts. However, those with pathogenic *PKD1* variants had a significantly earlier age of diagnosis (mean: 27.94 years vs 35.76 years; $p<0.05$) (**Supplementary Table 4**).

18% of our cohort had NMD. Patients with NMD were more likely to be male, older, had reduced baseline eGFR but also lower ΔeGFR and MKL compared to other genotype groups (**Supplementary Table 5**). They were less likely to have a known family history, more likely to have presented incidentally and had a lower prevalence of liver cysts (69% v 27%).

Of interest, we detected 29 recurrent variants in 91 unrelated pedigrees (**Supplementary Table 6**). The most common recurrent *PKD1* pathogenic variant, c.2534T>C, p.(Leu845Ser), was found in 11 different pedigrees. For *PKD2*, the most frequent pathogenic variant was c.2224C>T, p.(Arg742*), present in 6 pedigrees. 16 of the 29 variants were identified in UKBB but occurred with a different allele frequency (**Supplementary Figure 3**).

Genotype-phenotype correlations

61% of the cohort had a baseline eGFR >60ml/min/1.73m^2 (**Figure 2B**). Kaplan-Meier plots (probability of baseline eGFR>60ml/min/1.73m^2 v age) confirmed differences in median age between *PKD1-T*, *PKD2* and *NMD* patient groups (median age 42, 57 and 58 years respectively; $p>0.05$) (**Figure 4A**; **Supplementary Table 7**).

In patients with genotype, height-adjusted MKL and ΔeGFR information (n=95), *PKD1-T* patients showed the fastest rate of eGFR decline compared to *PKD1-NT*, *PKD-2*, *NMD* and "Others" (median: 4.6 ml/min/year vs. 2.2, 2.1, 2.2 and 0.8ml/min/year respectively; $p<0.05$) (**Figure 4B, Supplementary Figure 4A, B**). A visual representation of the average eGFR values across the different age groups (per decade) by genotype up to the 5th decade of life showed significant differences in mean eGFR per decade between *PKD1-T, PKD1-NT* and *PKD2* genotype (40.94 vs.

57.86, 61.35ml/min/1.73m^2 respectively (**Supplementary Figure 4C, Supplementary Table 8**). Nonetheless, in 3 genotyped pedigrees with at least 3 members with ΔeGFR values, there was significant intrafamilial variability in the rate of eGFR decline regardless of mutation type indicating a major influence of non-allelic factors in determining individual prognosis (**Supplementary Figure 4D, Supplementary Table 9**).

By linear regression, a significant positive correlation was observed between HtMKL and age for the *PKD1-T, PKD1-NT* and *PKD2* groups (**Figure 4C**, **Supplementary Table 10**). *PKD1-T, PKD1-NT and PKD2* genotypes had significantly higher median baseline HtMKL compared to NMD (8.9cm, 8.2cm, 8.3cm respectively vs 7.0cm, n=330) (**Figure 4D**).

Significant factors associated with rapid disease progression

The mean ΔeGFR for the cohort was -3.2 ± 2.3 ml/min/year. To assess the contribution of clinical factors apart from genotype to disease progression, the cohort was divided into 3 ΔeGFR groups (≤2.5 or slow, 2.5-5 or rapid, >5 or very rapid ml/min/year, *n* = 158) (**Table 2**). An earlier age at presentation, genotype (*PKD1-T*), proteinuria (uPCR >50mg/mmol), urinary tract infections (UTI) and higher baseline MKL were significant features of the very rapid ΔeGFR group compared to the slow ΔeGFR group.

By univariate logistic regression, ΔeGFR was significantly correlated with HtMKL >9.5 cm/m, MKL >16.5cm, a *PKD1-T* genotype, age at 1st clinic presentation ≤46 years and age of diagnosis ≤35 (**Supplementary Table 11**, n=70). We excluded genotypes other than *PKD1* and *PKD2* due to their low likelihood of developing kidney failure. Multivariate logistic regression confirmed that a *PKD1-T* genotype and HtMKL >9.5 cm/m were significant independent predictors of rapid disease progression (**Table 3**).

Development of a prognostic model to predict rapid disease progression

A prognostic model was derived from these two variables given equal weighting. The combination of both variables gave the highest positive predictive value (PPV, 88%) and specificity (89%) for rapid ΔeGFR (>2.5ml/min/year) compared to each factor alone (**Table 4**). Subgroup analyses based on baseline CKD stage and age at presentation showed that the model performed better in older patients (>40years) and those with more advanced disease (CKD3-4) (**Supplementary Tables 12-15**). In the subgroup of patients reaching kidney failure, the model predicted the onset of kidney failure

before age 60 years with a PPV of 100% and NPV of 63% (**Table 5, Supplementary Figure 2**). The model however performed less well (PPV 38%) in predicting very rapid progression (ΔeGFR >5ml/min/year) though better at excluding it (NPV 86%). This likely points to other unknown factors accelerating disease in this subgroup (**Supplementary Tables 16**).

In exploratory analysis, we compared the performance of our model with the MIC in a subgroup of patients with available HtTKV data and typical disease (**Supplementary Table 17,** n=34). Unexpectedly, our model showed a higher PPV (100% v 80%) and NPV (42% v 33%) compared to Mayo Class 1C-E (**Table 6**).

Discussion

In this paper, we report the clinical characteristics, presentation and genetic architecture of a longitudinal ADPKD cohort assessed and followed up for over a decade through a specialist clinic based in a regional kidney centre, serving a catchment population of 1.5 million through the UK National Health Service, a publicly funded system. As reflected in its age distribution (**Figure 2A**), this cohort is likely to be representative of the patient landscape within the UK population. By studying a cohort with relatively preserved kidney function longitudinally for a significant duration (mean follow-up 11.10 years), we were able to obtain important information regarding the natural history of ADPKD, including the 20% of patients who reached kidney failure during this period.

ADPKD patients presented in three major ways to our clinic: typical symptoms often abdominal or lumbar pain (44%), asymptomatic screening due to a positive family history (37%) or incidentally due to abdominal ultrasound scanning for other complaints (18%). Curiously, male patients presented at an older age and later disease stage than females: this may in part be due to the inclusion of older patients with later-onset atypical disease detected incidentally (**Supplementary Table 5**). Given the increasing use of imaging in routine diagnostics, it is likely that greater numbers of such patients will present incidentally.

The genetic architecture of our cohort yielded some unexpected results. First, the percentage of patients with pathogenic *PKD1* variants was only 53%. This figure likely reflects the much higher percentage of genetically unresolved cases compared to published cohorts: NMD in 18% and VUS in 6%, the latter largely related to *PKD1* [28-30]. The higher reporting rate of VUS likely relates to more stringent diagnostic criteria used in clinical diagnostics compared to research studies. The high rate of NMD could reflect the composition of clinical referrals rather than patient recruitment to selected research cohorts. We also included patients with late onset atypical presentations (some with more liver than kidney cysts) to capture new genetic variants. Nonetheless, our results are similar to those of a recent population-based study with electronic health record data which utilised exome sequencing (12 candidate genes analysed) [31]. Overall, 71.4% of those with pathogenic and likely pathogenic variants in *PKD1* or *PKD2* had a confirmed diagnosis of ADPKD. This difference was clearly due to the divergence between loss-of-function *PKD1* or *PKD2* variants (97-100% had ADPKD) and *PKD1* missense changes (31.2% had ADPKD), indicating variable penetrance in the latter or misclassification from VUS. Finally, it is worth noting that almost all of these patients were screened via a custom 17-gene cystic panel (Sheffield) [19] which did not include several of the newer 'ADPKD genes' ie *IFT140, ALG5, ALG8, ALG9* [7, 8, 32, 33]. The clinical characteristics of our NMD patients

nonetheless suggest a more benign course for disease progression and kidney failure, an important issue for patient management.

A small percentage of patients (6%) were identified with variants in other ADPKD genes such as *DNAJB11, GANAB, IFT140, PRKSCH, PKHD1* and ~ 1% had more complex genetic inheritance patterns (mosaic, biallelic, digenic). Although rare, identification of these 7% of patients is important for accurate diagnosis, prognosis, genetic counselling and decision-making relating to treatment eligibility. Finally, we discovered 91 unrelated pedigrees representing 29 different recurrent *PKD1* and *PKD2* variants in our cohort. Their allele frequency in UKBB however did not indicate that they are more widely distributed variants within the UK population.

The major finding of this study was the identification of two independent variables that predict ΔeGFR >2.5ml/min/1.73m^2/year ie a *PKD1-T* genotype and HtMKL >9.5cm/m. The combination of both factors allowed us to predict rapid disease progressors with a PPV of 88%. The model showed improved performance in older patients (>40 years) or more advanced CKD (stage 3-4) with an increase in the PPV to 100% and 95% respectively and predicted the onset of kidney failure before age 60 years with a PPV of 100%. Although TKV was not available in the majority of patients, we conducted an exploratory analysis in a subset of patients where TKV could be measured or estimated from clinical imaging. Unexpectedly, our model gave higher PPV (85% v 77%) and NPV (50% v 33%) readings compared to MIC risk groups (1C-E).

A previous report of intrafamilial variability including milder PKD (MIC 1A, B) in 18% of patients with a *PKD1-T* genotype demonstrates the importance of non-allelic factors in modifying the phenotype even in the highest risk group [34]. We confirm that this is independent of the familial genotype (**Supplementary Figure 4D**) and document the huge variation in ΔeGFR among family members in 3 pedigrees with at least 3 family members (**Supplementary Table 10**), similar findings to the age of kidney failure in other pedigrees [35]. Future research should concentrate on identifying other prognostic factors determining individual variability [36, 37] which could be included in a modified equation.

Our study has some limitations. First, although unselected, this was a single centre UK cohort study. Our experience may therefore not reflect those of other non-UK populations or differently funded health systems. Second, our patients were predominantly of White ethnicity and therefore our results may not apply to other ethnic groups. Third, routine US measurements were performed by a variety of clinical sonographers who did not undergo standardised training or assessment specific to the study. This is likely to have added to the variability of MKL measurements but conversely, removed bias and reflects real-world practice. Also, in deriving HtMKL, we were unable to exclude

patients with atypical disease (Class 2) from further analysis, unlike for HtTKV. If this was done, the sensitivity and specificity of HtMKL in predicting ΔeGFR might be further improved. Fourth, TKV measurements were obtained from retrospective historical imaging not optimised for MRI-TKV volumetry; in addition, CT-TKV values were derived from the ellipsoid equation. This could have reduced the sensitivity and specificity of the MIC applied in the sub-cohort analysed. Finally, we did not have available all the key age-dependent clinical variables needed to calculate individual PROPKD scores hence were unable to directly compare the performance of our model to the PROPKD score. It is interesting to note however that our PPV and NPV values are comparable to the PROPKD: a PROPKD score >6 (high risk group predicting kidney failure by age 60) had a PPV of 90.9% and NPV of 57.3% [14]. Our model appears to be at least equivalent in performance to PROPKD although this needs to be confirmed in a validation cohort where the performance of both could be directly compared.

Conclusion

Real-world longitudinal data confirms the importance of genotype and height-adjusted mean kidney length as independent variables determining ΔeGFR with improved performance shown in patients reaching kidney failure. The model described could improve patient selection for treatment.

Disclosure / Conflict of interest statement

ACMO is an Editor for NDT. The other authors declare no conflicts of interest. The results presented in this article have not been published previously in whole or part, except in abstract form.

Acknowledgements

We thank the patients and referring physicians for taking part in this study, Peter Metherall (3D Lab, Sheffield Teaching Hospitals) for assistance with TKV measurements. This project was conducted using summary statistics generated from the UK Biobank resource (application 26041 and 48511).

Authors' contributions

EWEC, JC, MKV, MD, RS, PC collected and/or analysed data; EWEC and ACMO wrote the paper; ACMO designed and supervised the study. All authors read and approved the final manuscript.

Funding

JC was supported by a Medical Research Council (MRC) Clinical Research Training Fellowship (MR/R01745X/1). RS was supported by a Clinical Lecturer starter grant from the Academy of Medical Sciences UK and a MRC Confidence in Concept award.

Data availability statement

All data are incorporated into the article and its online supplementary material.

Supplementary material

Detailed methods, Supplementary Figures 1-4, Supplementary Tables 1-16

Figure legends

Figure 1

Flowchart summarizing the analysis of 618 patients presenting to a specialist ADPKD clinic at Sheffield between 2010 and 2021. * indicates the total numbers available for each investigation.

Figure 2

A. Age distribution of patients in the cohort (n=618); B. Distribution of CKD classes in the total cohort (n=609);

Figure 3

A. Distribution of different genotypes in this cohort (n=549); B. Distribution of VUS detected in different genotypes (n=38).

Figure 4

A. Kaplan Meier plot (probability of eGFR >60ml/min/1.73m^2) for different genotypes with age (n=55); B. Correlation of genotype with the rate of eGFR decline in patients with HtMKL and genotype data (n=95); C. Relationship between HtMKL and age by genotype (n=367); D. Relationship of genotype with HtMKL (cm) (n=330).

Tables

Table 1: Demographics and clinical characteristics of patients in the cohort

	***n* = 618**
Age of 1st clinic assessment, years	45.66 ± 17.63
Age of diagnosis, years	37.16 ± 19.59
Follow up duration, years	11.11 ± 11.20
Gender	
Male, *n* (%)	309 (50%)
Female, *n* (%)	309 (50%)
Ethnicity	
White, *n* (%)	581 (94%)
Asian (Indian, Pakistani, Bangladeshi, Chinese), *n* (%)	23 (4%)
Other ethnic groups (Arab, any other ethnic group), *n* (%)	5 (1%)
Black (African, Caribbean), *n* (%)	5 (1%)
Mixed ethnic group (any mixed or multiple ethnicity), *n* (%)	4 (1%)
Family history of ADPKD, *n* (%)	452 (73%)
Comorbidity	
Hypertension at presentation, *n* (%)	402 (65%)
Hernia, *n* (%)	83 (13%)
Gout, *n* (%)	65 (10%)
Diabetes, *n* (%)	38 (6%)
Types of initial presentation at diagnosis	
Symptomatic, *n* (%)	275 (44%)
Lumbar/abdominal pain, *n* (%)	110 (18%)
Hypertension, *n* (%)	54 (9%)
UTI, *n* (%)	41 (7%)
CKD, *n* (%)	33 (5%)
Haematuria, *n* (%)	25 (4%)
Abdominal mass, *n* (%)	12 (2%)
Screening, *n* (%)	231 (37%)
Incidental, *n* (%)	112 (18%)
Imaging (including antenatal scans), *n* (%)	103 (17%)
Abnormal biochemistry (U&E, LFTs), *n* (%)	9 (1%)
Renal manifestations	
UTI, *n* (%)	201 (33%)
Lumbar/abdominal pain, *n* (%)	153 (25%)
Macroscopic haematuria, *n* (%)	103 (17%)

Significant proteinuria (uPCR >50mg/mmol) (*n* = 568), *n* (%)	50 (9%)
Renal calculi, *n* (%)	43 (7%)
Smoking	
Non-smoker, *n* (%)	335 (54%)
Ex-smoker, *n* (%)	191 (31%)
Active smoker, *n* (%)	92 (15%)
Height	
Mean height (n = 518), m	1.73 ± 0.10
Blood pressure	
Mean systolic BP at presentation (*n* = 608), mm Hg	144.11 ± 15.98
Mean diastolic BP at presentation (*n* = 608), mm Hg	82.25 ± 10.67
Number of antihypertensives for those with hypertension (*n* = 401)	
Zero, *n* (%)	81 (20%)
One, *n* (%)	124 (31%)
Two, *n* (%)	122 (30%)
Three, *n* (%)	67 (17%)
Four, *n* (%)	8 (2%)

Laboratory findings	
Baseline creatinine (*n* = 609), μmol/L	117.86 ± 77.40
Baseline eGFR (*n* = 609), ml/min/1.73m^2	74.27 ± 35.61
Follow up creatinine (*n* = 578), μmol/L	142.44 ± 116.34
Follow up eGFR (*n* = 578), ml/min/1.73m^2	67.36 ± 37.00
ΔGFR change (*n* = 188), ml/min/year	-3.18 ± 2.27

Data are mean and standard deviation, median and interquartile range, or as absolute number and % frequency, as appropriate.

Table 2: Comparing clinical characteristics of patients with slow, rapid and very rapid rates of eGFR decline (n = 158)

Rate of eGFR decline, ml/min/1.73m^2/year	≤2.5 (*n* = 75) (Slow rate)	2.5-5 (*n* = 53) (Rapid rate)	>5 (*n* = 30) (Very rapid rate)	p-value (<2.5 vs >5)	p-value (<2.5 vs 2.5-5)	p-value (2.5-5 vs >5)
Age of 1st clinic assessment, year	59.79 ± 12.33	52.92 ± 13.14	**46.80 ± 10.27**	**<0.0001***	**0.0035***	**0.0213***
Age of diagnosis, year	48.43 ± 16.09	36.89 ± 16.91	35.60 ± 14.75	**<0.0002***	**0.0002***	0.7187
Follow up duration**,** years	15.55 ± 11.18	20.45 ± 13.14	15.80 ± 10.52	0.9132	**0.0293***	0.0818
Gender						
Male, *n* (%)	45 (60%)	30 (57%)	15 (50%)	0.3496	0.7008	0.5618
Female, *n* (%)	30 (40%)	23 (43%)	15 (50%)	0.3496	0.7008	0.5618
Family history of ADPKD, *n* (%)	55 (73%)	44 (83%)	23 (77%)	0.7241	0.1973	0.4809
Genotypes						
PKD1-T, *n* (%)	**5 (7%)**	**20 (38%)**	**20 (67%)**	**<0.0001***	**<0.0001***	**0.0113***
PKD1-NT, *n* (%)	22 (29%)	15 (32%)	5 (17%)	0.1797	0.8991	0.2337
PKD2, *n* (%)	22 (29%)	11 (21%)	2 (7%)	0.0125	0.2744	0.0898
No mutation detected (NMD), *n* (%)	20 (27%)	6 (11%)	3 (10%)	0.0621	**0.0335***	0.8525
Others (DNAJB11, PKHD1, GANAB, IFT140), *n* (%)	6 (8%)	1 (2%)	0 (0%)	0.1106	0.1341	-
Comorbidity						
Hypertension at presentation, *n* (%)	67 (89%)	46 (87%)	26 (87%)	0.698	0.6598	0.987
Gout, *n* (%)	15 (20%)	8 (15%)	3 (10%)	0.2193	0.4764	0.5108
Hernia, *n* (%)	6 (8%)	11 (21%)	4 (13%)	0.4003	**0.0362***	0.3986
Diabetes, *n* (%)	2 (3%)	3 (6%)	3 (10%)	0.1109	0.3892	0.4633
Types of initial presentation at diagnosis						
Symptomatic, *n* (%)	**34 (45%)**	**28 (53%)**	**20 (67%)**	**0.0482***	0.4032	0.2201
Screening, *n* (%)	23 (31%)	18 (34%)	10 (33%)	0.7903	0.6939	0.9536
Incidental, *n* (%)	**18 (24%)**	**7 (13%)**	**0 (0%)**	**0.0032***	0.1292	**0.0375***
Renal manifestations						
Macroscopic haematuria, *n* (%)	13 (17%)	14 (26%)	6 (20%)	0.7485	0.2148	0.5115
UTI, *n* (%)	**21 (28%)**	**18 (34%)**	**15 (50%)**	**0.0319***	0.4704	0.1515
Lumbar/abdominal pain, *n* (%)	24 (32%)	17 (32%)	7 (23%)	0.3791	0.9928	0.3987
Renal calculi, *n* (%)	7 (9%)	4 (8%)	3 (10%)	0.9163	0.7725	0.6993
Significant proteinuria (uPCR >50mg/mmol) (*n* = 156), *n* (%)	**5 out of 74 (7%)**	**2 out of 52 (4%)**	**7 out of 30 (23%)**	**0.0165***	0.4826	**0.0065***

Smoking						
Active smoker, *n* (%)	12 (16%)	9 (17%)	2 (7%)	0.2037	0.8826	0.183
Ex-smoker, *n* (%)	23 (31%)	16 (30%)	10 (33%)	0.7903	0.9539	0.7667
Non-smoker, *n* (%)	40 (53%)	28 (53%)	18 (60%)	0.5348	0.9552	0.5278
Baseline eGFR, ml/min/1.73m^2	52.33 ± 22.74 (*n = 75*)	49.75 ± 22.62 (n=53)	54.64 ± 21.41 (n=30)	0.6256	0.5277	0.3314
Follow up eGFR, ml/min/1.73m^2	**44.21 ± 21.09 (*n = 74*)**	**34.67 ± 19.27 (n=51)**	**30.06 ± 18.49 (*n = 29*)**	**0.0014***	**0.0102***	0.2946
Baseline height-adjusted mean kidney length on USS (*n* = 162), cm	**8.95 ± 2.19 (*n = 55*)**	**9.85 ± 1.77 (*n* = 43)**	**10.68 ± 2.40 (*n = 22*)**	**0.0059***	**0.0274***	0.1602
Mean number of anti-hypertensives at baseline, types	1.67 ± 1.02	1.91 ± 1.18	1.57 +/- 1.01	0.6484	0.2356	0.1711

Data are mean and standard deviation, median and interquartile range, or as absolute number and % frequency, as appropriate.

Table 3: Multivariate logistic regression analysis (n = 70)

Variable	*Reference category*	*ΔeGFR >2.5ml/min/year**		*P-value*
		Odds Ratio	*95% CI*	
Genotype				
PKD1(T)	*PKD2*	5.980	1.51 to 27.4	0.014**
PKD1(NT)		0.970	0.26 to 3.63	0.964
Height adjusted Mean Kidney Length (HtMKL) baseline, cm/m				
HtMKL >9.5 cm/m	HtMKL ⩽9.5 cm/m	3.228	1.02 to 10.54	0.047**

*C statistic for Δ*eGFR >2.5ml/min/year* model = 0.7851, p<0.0001

** P<0.05

Table 4: Predictive values of baseline factors on the likelihood of rapid disease progression (ΔGFR >2.5ml/min/year)

n=70	*PKD1-T*	HtMKL >9.5 cm/m	*PKD1-T* + HtMKL >9.5cm/m
PPV	87%	76%	88%
NPV	58%	64%	55%
Specificity	85%	59%	89%
Sensitivity	60%	79%	53%

TP = true positive; TN = true negative; FP = false positive; FN = false negative

Sensitivity = TP/(TP+FN); **Specificity** = TN/(TN+FP); **Positive predictive value (PPV)** = TP/(TP+FP);

Negative predictive value (NPV) = TN/(TN+FN)

Table 5: Predictive values of baseline factors on the likelihood of kidney failure by age 60 years

n=34	*PKD1-T*	HtMKL >9.5 cm/m	*PKD1-T* + HtMKL >9.5cm/m
PPV	95%	81%	100%
NPV	64%	83%	63%
Specificity	90%	50%	100%
Sensitivity	78%	96%	74%

Table 6: Predictive values of baseline factors on the likelihood of rapid disease progression compared to MIC

n=34	MIC (1C, 1D, 1E)	*PKD1-T*	HtMKL >9.5 cm/m	*PKD1-T* + HtMKL >9.5cm/m
PPV	80%	94%	83%	100%
NPV	33%	44%	40%	42%
Specificity	38%	88%	50%	100%
Sensitivity	77%	65%	77%	58%

References

1. Ong AC, Devuyst O, Knebelmann B*, et al.* Autosomal dominant polycystic kidney disease: the changing face of clinical management. Lancet 2015;385(9981):1993-2002
2. Willey CJ, Blais JD, Hall AK*, et al.* Prevalence of autosomal dominant polycystic kidney disease in the European Union. Nephrol Dial Transplant 2016
3. Lanktree MB, Haghighi A, Guiard E*, et al.* Prevalence Estimates of Polycystic Kidney and Liver Disease by Population Sequencing. J Am Soc Nephrol 2018;29(10):2593-2600
4. Dalgaard OZ. Bilateral polycystic disease of the kidneys; a follow-up of two hundred and eighty-four patients and their families. Acta Med Scand 1957;158(Suppl 328):1-255
5. Ong AC, Harris PC. A polycystin-centric view of cyst formation and disease: the polycystins revisited. Kidney Int 2015;88(4):699-710
6. Cornec-Le Gall E, Torres VE, Harris PC. Genetic Complexity of Autosomal Dominant Polycystic Kidney and Liver Diseases. J Am Soc Nephrol 2018;29(1):13-23
7. Besse W, Chang AR, Luo JZ*, et al.* ALG9 Mutation Carriers Develop Kidney and Liver Cysts. J Am Soc Nephrol 2019;30(11):2091-2102
8. Apple B, Sartori G, Moore B*, et al.* Individuals heterozygous for ALG8 protein-truncating variants are at increased risk of a mild cystic kidney disease. Kidney Int 2023;103(3):607-615
9. Ong ACM. Polycystic kidney disease: Tolvaptan slows disease progression in late-stage ADPKD. Nat Rev Nephrol 2018
10. Torres VE, Gansevoort RT, Czerwiec FS. Tolvaptan in Later-Stage Polycystic Kidney Disease. N Engl J Med 2018;378(5):489-490
11. Gansevoort RT, Arici M, Benzing T*, et al.* Recommendations for the use of tolvaptan in autosomal dominant polycystic kidney disease: a position statement on behalf of the ERA-EDTA Working Groups on Inherited Kidney Disorders and European Renal Best Practice. Nephrol Dial Transplant 2016;31(3):337-348
12. Chong J, Harris T, Ong ACM. Regional variation in tolvaptan prescribing across England: national data and retrospective evaluation from an expert centre. Clin Kidney J 2023;16(1):61-68
13. Irazabal MV, Rangel LJ, Bergstralh EJ*, et al.* Imaging classification of autosomal dominant polycystic kidney disease: a simple model for selecting patients for clinical trials. J Am Soc Nephrol 2015;26(1):160-172
14. Cornec-Le Gall E, Audrezet MP, Rousseau A*, et al.* The PROPKD Score: A New Algorithm to Predict Renal Survival in Autosomal Dominant Polycystic Kidney Disease. J Am Soc Nephrol 2016;27(3):942-951
15. Bhutani H, Smith V, Rahbari-Oskoui F*, et al.* A comparison of ultrasound and magnetic resonance imaging shows that kidney length predicts chronic kidney disease in autosomal dominant polycystic kidney disease. Kidney Int 2015;88(1):146-151
16. Barua M, Cil O, Paterson AD*, et al.* Family history of renal disease severity predicts the mutated gene in ADPKD. J Am Soc Nephrol 2009;20(8):1833-1838
17. Chapman AB, Bost JE, Torres VE*, et al.* Kidney volume and functional outcomes in autosomal dominant polycystic kidney disease. Clinical journal of the American Society of Nephrology : CJASN 2012;7(3):479-486
18. Pei Y, Hwang YH, Conklin J*, et al.* Imaging-based diagnosis of autosomal dominant polycystic kidney disease. J Am Soc Nephrol 2015;26(3):746-753
19. Durkie M, Chong J, Valluru MK*, et al.* Biallelic inheritance of hypomorphic PKD1 variants is highly prevalent in very early onset polycystic kidney disease. Genet Med 2021;23(4):689-697
20. Richards S, Aziz N, Bale S*, et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17(5):405-424
21. Matsushita K, Mahmoodi BK, Woodward M*, et al.* Comparison of risk prediction using the CKD-EPI equation and the MDRD study equation for estimated glomerular filtration rate. JAMA : the journal of the American Medical Association 2012;307(18):1941-1951
22. Gabow PA, Johnson AM, Kaehny WD*, et al.* Factors affecting the progression of renal disease in autosomal-dominant polycystic kidney disease. Kidney Int 1992;41(5):1311-1319
23. Chapman AB, Johnson AM, Rainguet S*, et al.* Left ventricular hypertrophy in autosomal dominant polycystic kidney disease. J Am Soc Nephrol 1997;8(8):1292-1297
24. Martinez V, Furlano M, Sans L*, et al.* Autosomal dominant polycystic kidney disease in young adults. Clin Kidney J 2023;16(6):985-995

25. Kelleher CL, McFann KK, Johnson AM*, et al.* Characteristics of hypertension in young adults with autosomal dominant polycystic kidney disease compared with the general U.S. population. Am J Hypertens 2004;17(11 Pt 1):1029-1034
26. Schrier RW, Brosnahan G, Cadnapaphornchai MA*, et al.* Predictors of Autosomal Dominant Polycystic Kidney Disease Progression. J Am Soc Nephrol 2014
27. Cadnapaphornchai MA, Ong ACM. Hypertension in young adults with autosomal dominant polycystic kidney disease: a case for early screening? Clin Kidney J 2023;16(6):901-904
28. Rossetti S, Consugar MB, Chapman AB*, et al.* Comprehensive molecular diagnostics in autosomal dominant polycystic kidney disease. J Am Soc Nephrol 2007;18(7):2143-2160
29. Hwang YH, Conklin J, Chan W*, et al.* Refining Genotype-Phenotype Correlation in Autosomal Dominant Polycystic Kidney Disease. J Am Soc Nephrol 2015
30. Audrezet MP, Cornec-Le Gall E, Chen JM*, et al.* Autosomal dominant polycystic kidney disease: comprehensive mutation analysis of PKD1 and PKD2 in 700 unrelated patients. Hum Mutat 2012;33(8):1239-1250
31. Chang AR, Moore BS, Luo JZ*, et al.* Exome Sequencing of a Clinical Population for Autosomal Dominant Polycystic Kidney Disease. JAMA 2022;328(24):2412-2421
32. Senum SR, Li YSM, Benson KA*, et al.* Monoallelic IFT140 pathogenic variants are an important cause of the autosomal dominant polycystic kidney-spectrum phenotype. Am J Hum Genet 2022;109(1):136-156
33. Lemoine H, Raud L, Foulquier F*, et al.* Monoallelic pathogenic ALG5 variants cause atypical polycystic kidney disease and interstitial fibrosis. Am J Hum Genet 2022;109(8):1484-1499
34. Lanktree MB, Guiard E, Akbari P*, et al.* Patients with Protein-Truncating PKD1 Mutations and Mild ADPKD. Clin J Am Soc Nephrol 2021;16(3):374-383
35. Thong KM, Ong AC. The natural history of autosomal dominant polycystic kidney disease: 30-year experience from a single centre. QJM : monthly journal of the Association of Physicians 2013;106(7):639-646
36. Magayr TA, Song X, Streets AJ*, et al.* Global microRNA profiling in human urinary exosomes reveals novel disease biomarkers and cellular pathways for autosomal dominant polycystic kidney disease. Kidney Int 2020;98(2):420-435
37. Streets AJ, Magayr TA, Huang L*, et al.* Parallel microarray profiling identifies ErbB4 as a determinant of cyst growth in ADPKD and a prognostic biomarker for disease progression. Am J Physiol Renal Physiol 2017;312(4):F577-F588

Figure 1

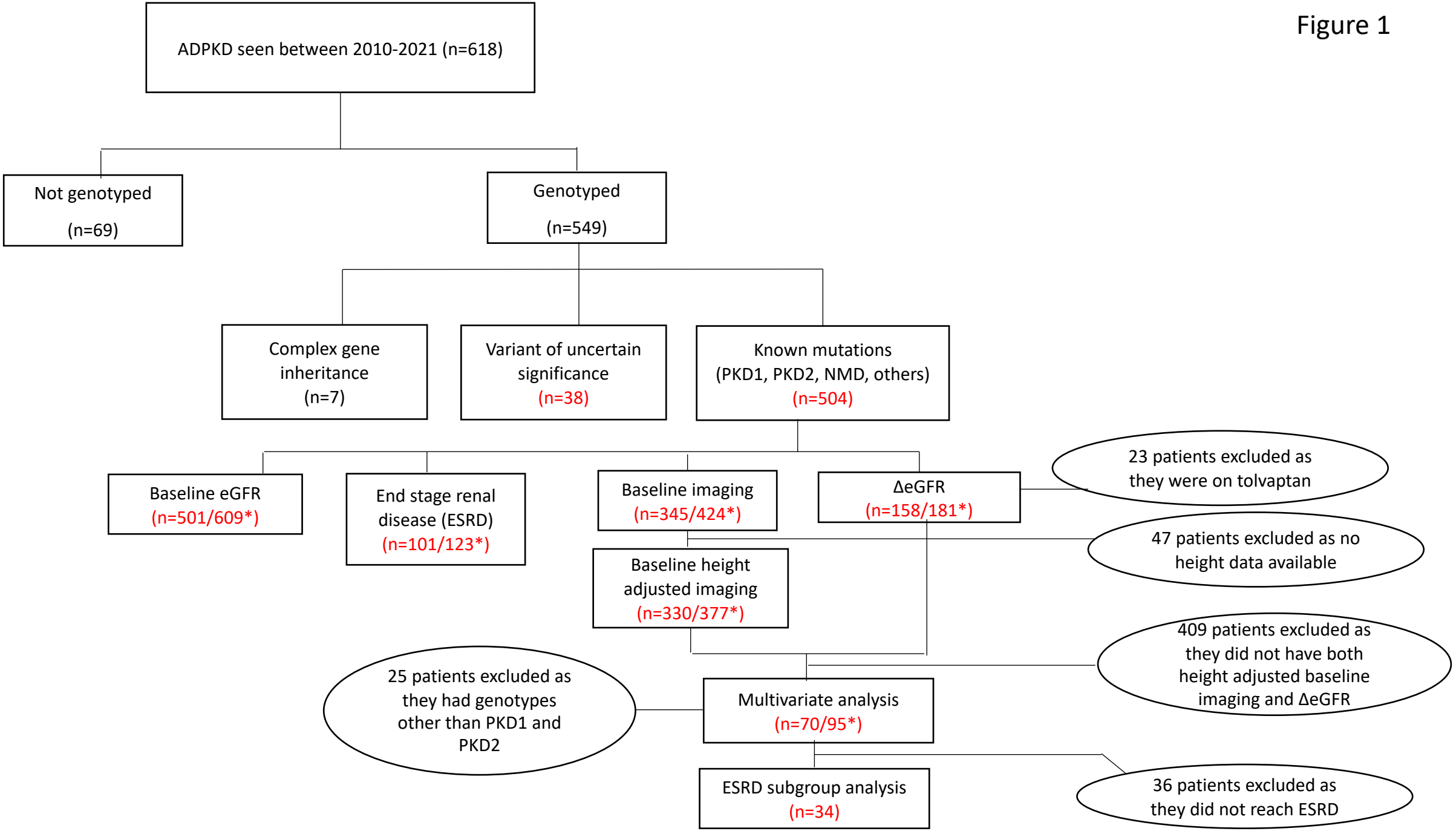
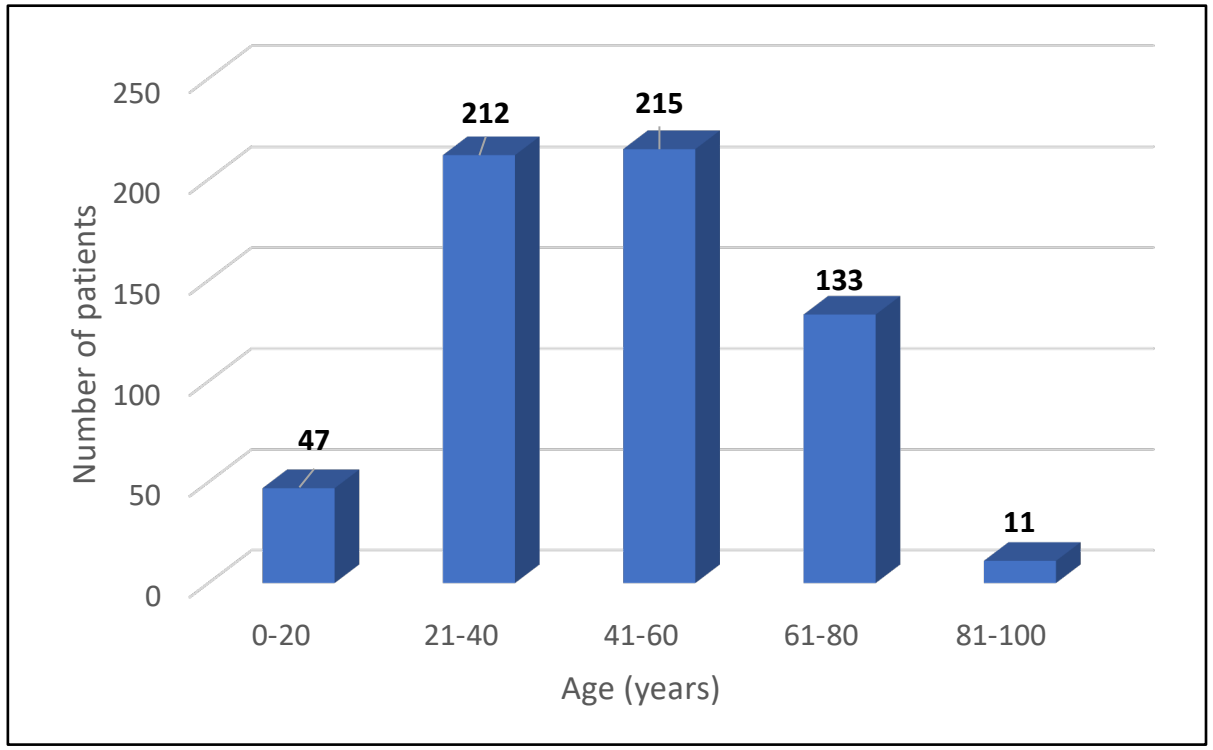

ADPKD seen between 2010-2021 (n=618)
Not genotyped (n=69)
Genotyped (n=549)
Complex gene inheritance (n=7)
Variant of uncertain significance (n=38)
Known mutations (PKD1, PKD2, NMD, others) (n=504)
Baseline eGFR (n=501/609*)
End stage renal disease (ESRD) (n=101/123*)
Baseline imaging (n=345/424*)
ΔeGFR (n=158/181*)
23 patients excluded as they were on tolvaptan
47 patients excluded as no height data available
Baseline height adjusted imaging (n=330/377*)
409 patients excluded as they did not have both height adjusted baseline imaging and ΔeGFR
25 patients excluded as they had genotypes other than PKD1 and PKD2
Multivariate analysis (n=70/95*)
ESRD subgroup analysis (n=34)
36 patients excluded as they did not reach ESRD

Figure 2

A


250
200
150
100
50
0
Number of patients
47
212
215
133
11
0-20
21-40
41-60
61-80
81-100
Age (years)

B

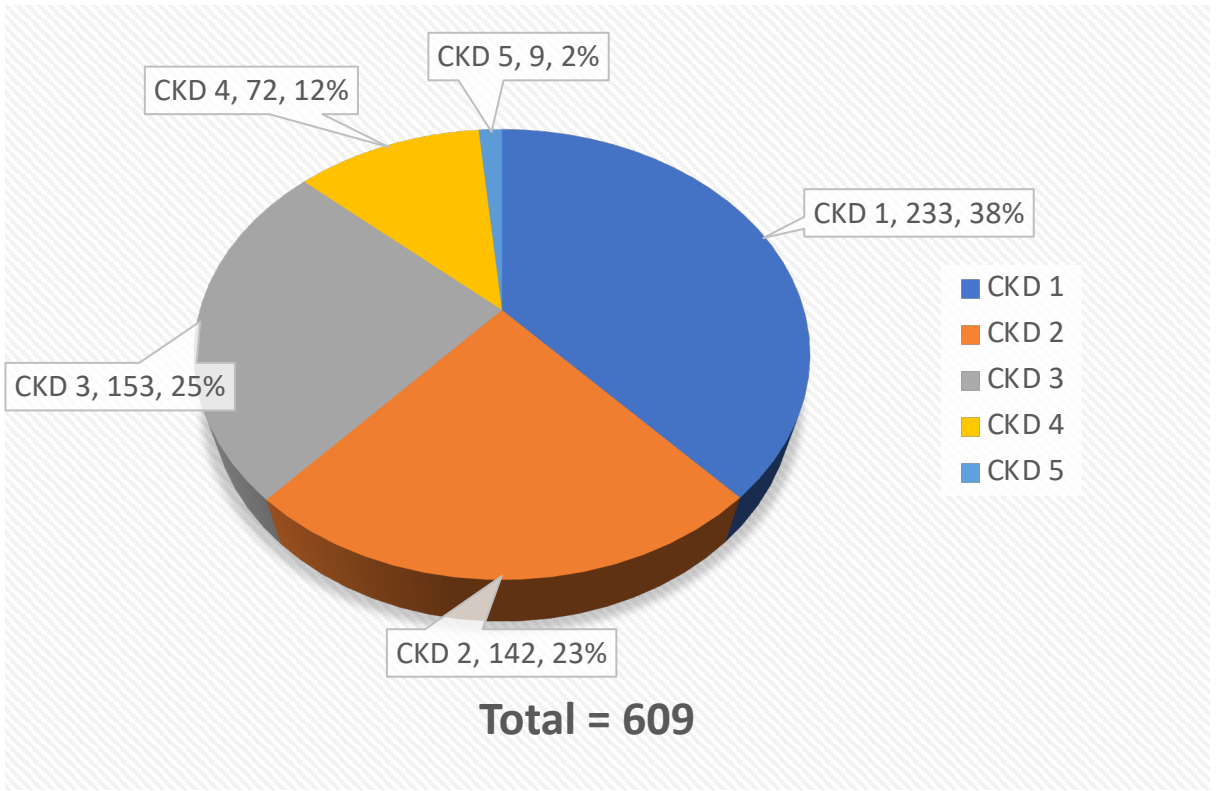
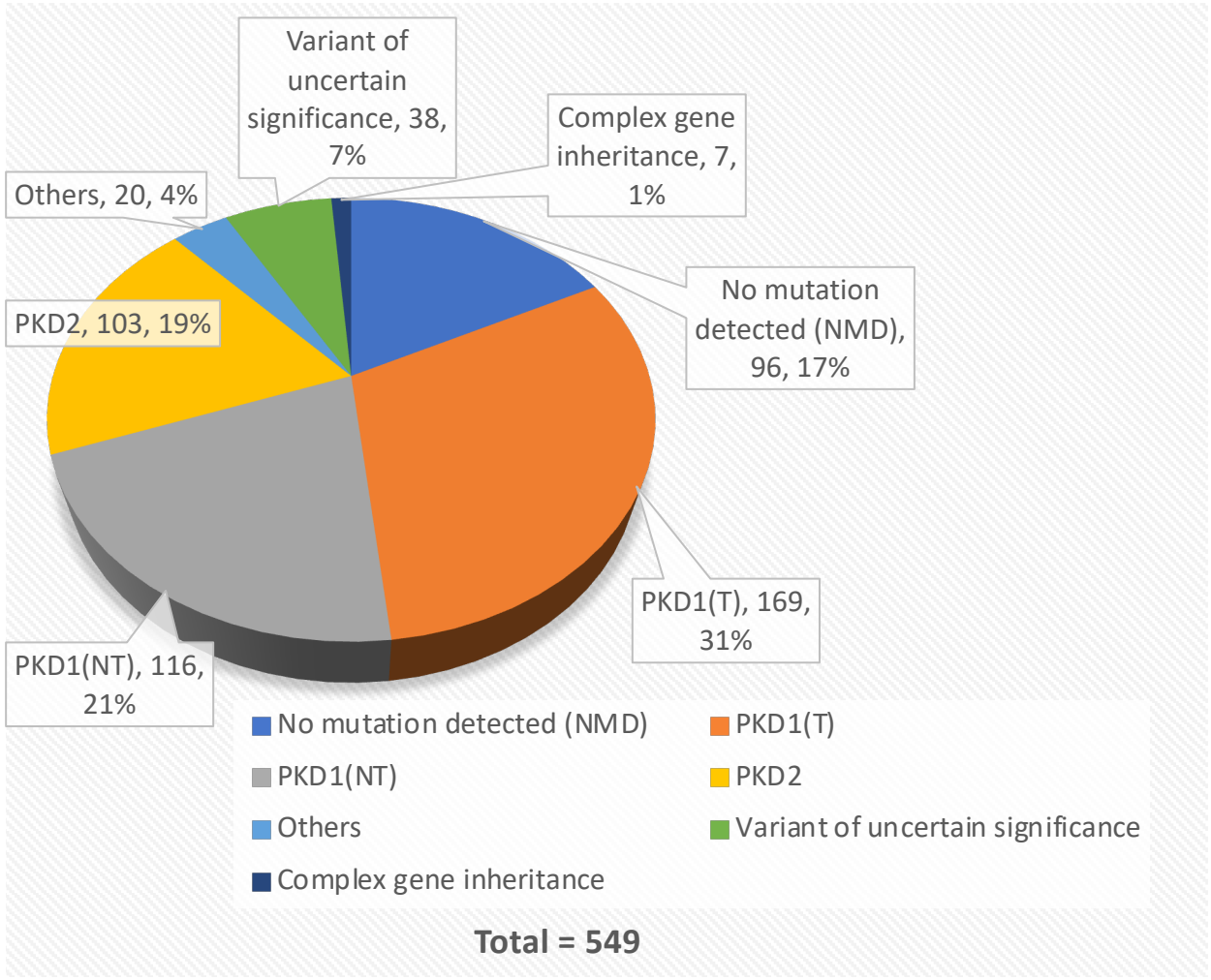

CKD 5, 9, 2%
CKD 4, 72, 12%
CKD 1, 233, 38%
CKD 3, 153, 25%
CKD 2, 142, 23%
CKD 1
CKD 2
CKD 3
CKD 4
CKD 5
Total = 609

Figure 3

A


Variant of uncertain significance, 38, 7%
Complex gene inheritance, 7, 1%
Others, 20, 4%
No mutation detected (NMD), 96, 17%
PKD2, 103, 19%
PKD1(T), 169, 31%
PKD1(NT), 116, 21%
No mutation detected (NMD)
PKD1(T)
PKD1(NT)
PKD2
Others
Variant of uncertain significance
Complex gene inheritance
Total = 549

B

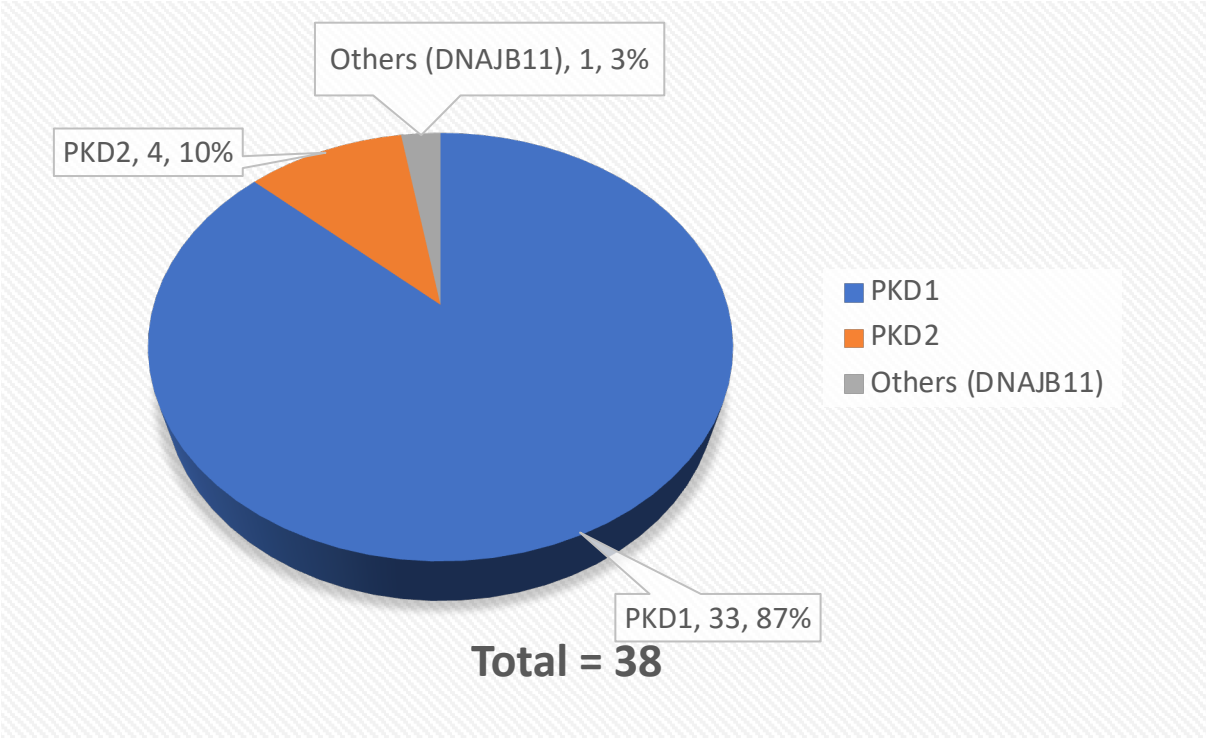
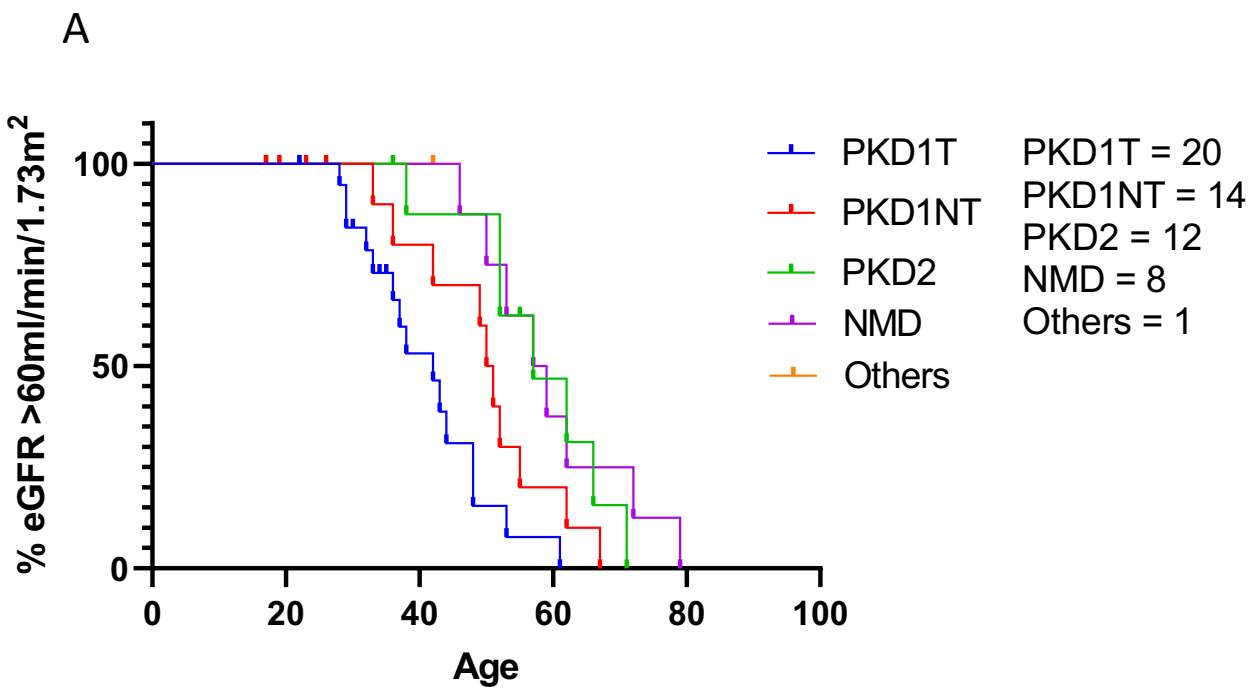
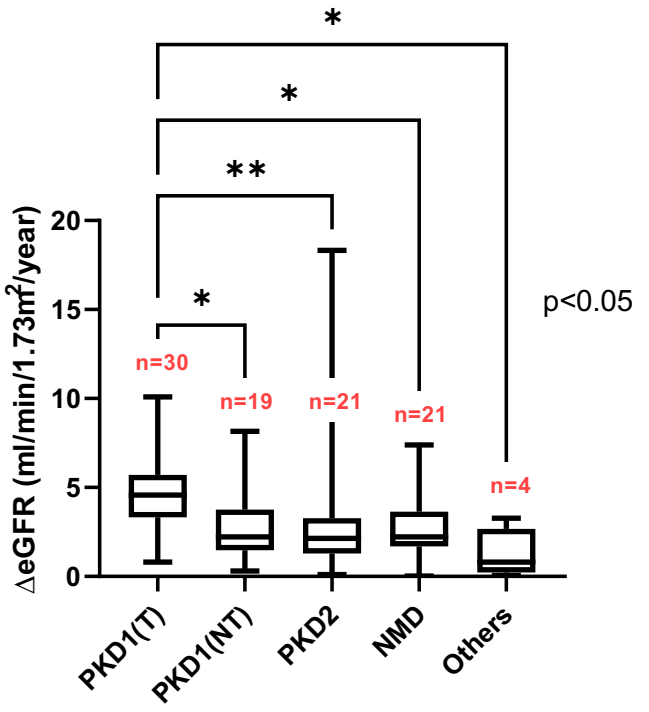
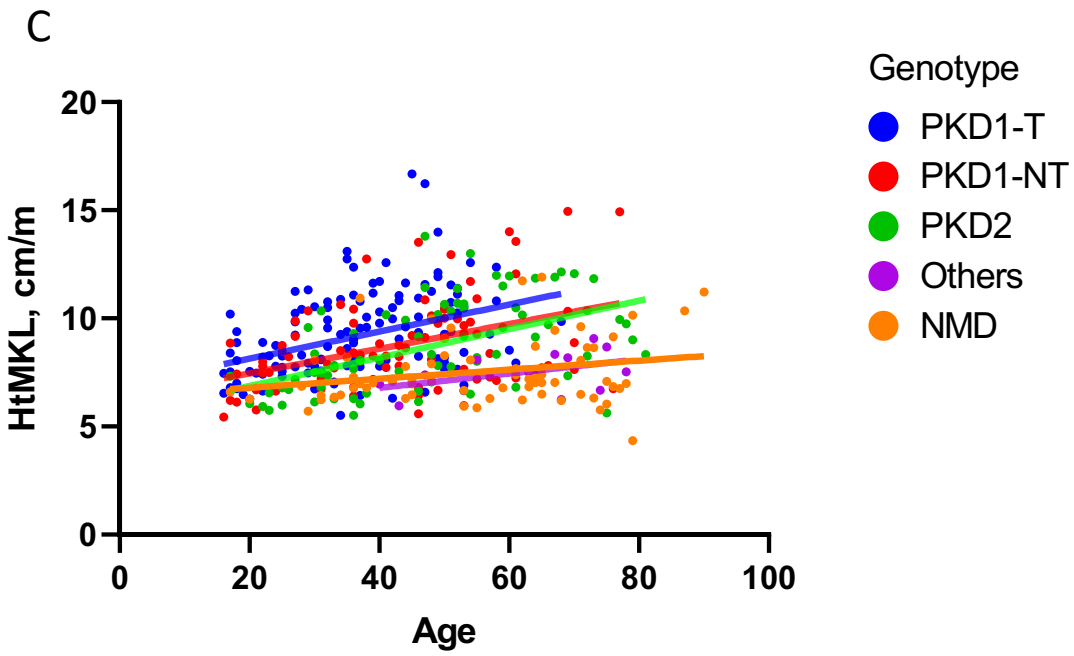
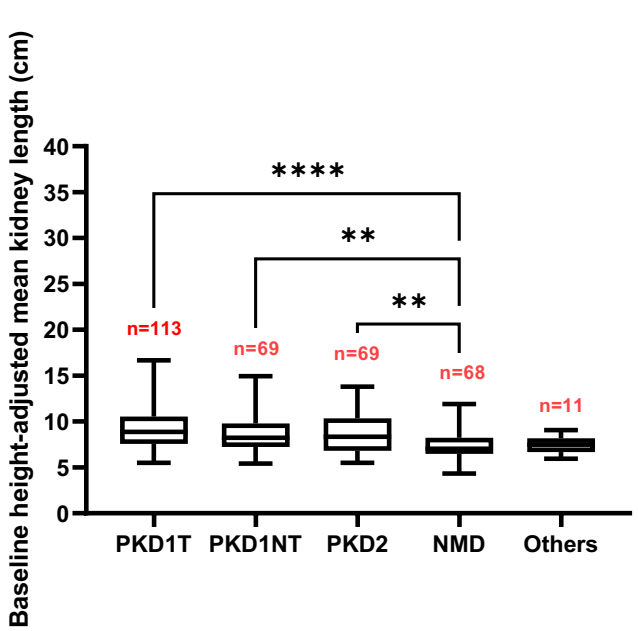

Others (DNAJB11), 1, 3%
PKD2, 4, 10%
PKD1, 33, 87%
PKD1
PKD2
Others (DNAJB11)
Total = 38

Figure 4


A
% eGFR >60ml/min/1.73m2
100
50
0
0
20
40
60
80
100
Age
PKD1T
PKD1NT
PKD2
NMD
Others
PKD1T = 20
PKD1NT = 14
PKD2 = 12
NMD = 8
Others = 1


B
*
*
**
*
ΔeGFR (ml/min/1.73m2/year)
20
15
10
5
0
n=30
n=19
n=21
n=21
n=4
p<0.05
PKD1(T)
PKD1(NT)
PKD2
NMD
Others


C
HtMKL, cm/m
20
15
10
5
0
0
20
40
60
80
100
Age
Genotype
PKD1-T
PKD1-NT
PKD2
Others
NMD


D
Baseline height-adjusted mean kidney length (cm)
40
35
30
25
20
15
10
5
0
****
**
**
n=113
n=69
n=69
n=68
n=11
PKD1T
PKD1NT
PKD2
NMD
Others