AJKD
Original Investigation

Estimation of Total Kidney Volume in Autosomal Dominant Polycystic Kidney Disease

Edwin M. Spithoven, MD, PhD,[1] Maatje D.A. van Gastel, BSc,[1,*] A. Lianne Messchendorp, MD,[1,*] Niek F. Casteleijn, MD,[1] Joost P.H. Drenth, MD, PhD,[2] Carlo A. Gaillard, MD, PhD,[1] Johan W. de Fijter, MD, PhD,[3] Esther Meijer, MD, PhD,[1] Dorien J.M. Peters, PhD,[4] Peter Kappert, MSc,[5] Remco J. Renken, PhD,[6] Folkert W. Visser, MD, PhD,[1] Jack F.M. Wetzels, MD, PhD,[7] Robert Zietse, MD, PhD,[8] and Ron T. Gansevoort, MD, PhD,[1] on behalf of the DIPAK Consortium[†]

Background: In autosomal dominant polycystic kidney disease (ADPKD), obtaining measured total kidney volume (mTKV) by magnetic resonance (MR) imaging and manual tracing is time consuming. Two alternative MR imaging methods have recently been proposed to estimate TKV ($eTKV_{ellipsoid}$ and $eTKV_{PANK}$), which require less time.

Study Design: Cross-sectional and longitudinal diagnostic test study.

Setting & Participants: Patients with ADPKD with a wide range of kidney function and an approved T2-weighted MR image obtained at the University Medical Centers of Groningen, Leiden, Nijmegen, and Rotterdam, the Netherlands, in 2007 to 2014. Test set for assessing reproducibility, n = 10; cohort for cross-sectional analyses, n = 220; and cohort for longitudinal analyses, n = 48.

Index Tests: Average times for $eTKV_{ellipsoid}$ and $eTKV_{PANK}$ were 5 and 15 minutes, respectively. Bias is defined as (mTKV − eTKV)/mTKV × 100%; precision, as one standard deviation of bias.

Reference Tests: mTKV using manual tracing to calculate the area within kidney boundaries times slice thickness. Average time for mTKV was 55 minutes.

Results: In the test set, intra- and intercoefficients of variation for mTKV, $eTKV_{ellipsoid}$, and $eTKV_{PANK}$ were 1.8% and 2.3%, 3.9% and 6.3%, and 3.0% and 3.4%, respectively. In cross-sectional analysis, baseline mTKV, $eTKV_{ellipsoid}$, and $eTKV_{PANK}$ were 1.96 (IQR, 1.28-2.82), 1.93 (IQR, 1.25-2.82), and 1.81 (IQR, 1.17-2.62) L, respectively. In cross-sectional analysis, bias was 0.02% ± 3.2%, 1.4% ± 9.2%, and 4.6% ± 7.6% for repeat mTKV, $eTKV_{ellipsoid}$, and $eTKV_{PANK}$, respectively. In longitudinal analysis, no significant differences were observed between percentage change in mTKV (16.7% ± 17.1%) and percentage change in $eTKV_{ellipsoid}$ (19.3% ± 16.1%) and $eTKV_{PANK}$ (17.8% ± 16.1%) over 3 years.

Limitations: Results for follow-up data should be interpreted with caution because of the limited number of patients.

Conclusions: Both methods for eTKV perform relatively well compared to mTKV and can detect change in TKV over time. Because $eTKV_{ellipsoid}$ requires less time than $eTKV_{PANK}$, we suggest that this method may be preferable in clinical care.

Am J Kidney Dis. 66(5):792-801.

INDEX WORDS: Autosomal dominant polycystic kidney disease (ADPKD); total kidney volume (TKV); magnetic resonance imaging (MRI); estimation methods; ellipsoid; PANK; validation.

Autosomal dominant polycystic kidney disease (ADPKD) is characterized by the formation and growth of numerous cysts in both kidneys, leading to an increase in kidney volume. These cysts compress healthy kidney tissue, causing progressive kidney function decline and, in most patients, ultimately a need for renal replacement therapy. In patients with ADPKD, total kidney volume (TKV) has been shown to be an early marker of disease severity and predictor of kidney function decline.[1] Measurement of TKV is therefore used to assess prognosis in clinical care and for selection of patients for randomized controlled

From the [1]Department of Nephrology, University Medical Center Groningen, Groningen; [2]Department of Gastroenterology and Hepatology, Radboud University Medical Center, Nijmegen; Departments of [3]Nephrology and [4]Human Genetics, Leiden University Medical Center, Leiden; [5]Department of Radiology, University Medical Center Groningen; [6]Neuroimaging Center, University of Groningen, University Medical Center Groningen, Groningen; [7]Department of Nephrology, Radboud University Medical Center, Nijmegen; and [8]Department of Nephrology, Erasmus Medical Center, Rotterdam, the Netherlands.

[*]MDAvG and ALM contributed equally to this work.

[†]A list of DIPAK Consortium members appears in the Acknowledgements.

Received October 2, 2014. Accepted in revised form June 8, 2015. Originally published online July 30, 2015.

Address correspondence to Ron T. Gansevoort, MD, PhD, Department of Nephrology, University Medical Center Groningen, University of Groningen, PO Box 30.001, 9700 RB Groningen, the Netherlands. E-mail: r.t.gansevoort@umcg.nl

0272-6386
http://dx.doi.org/10.1053/j.ajkd.2015.06.017

trials.[2] In these trials that investigate potential treatments for patients with ADPKD, assessment of TKV is often used as the primary or secondary study end point.[3-5]

The true gold-standard method to assess TKV is the manual tracing method. Computer tomogram or magnetic resonance (MR) images are used, and in each slice, the kidney boundaries are traced manually using dedicated software. Measured TKV (mTKV) is calculated from a set of contiguous images by summing the products of the area measurements within the kidney boundaries and slice thickness.[6] This method is laborious, which limits its use in trial settings, but especially in clinical care.

If kidney volume could be estimated with sufficient accuracy and reliability, it would alleviate the time-consuming process of kidney volume measurement. Recently, 2 kidney volume estimation methods have been developed: the midslice method[7] by the Consortium for Radiologic Imaging Studies of ADPKD (CRISP) and the ellipsoid method[2] by the Mayo Clinic. For both methods, measured and estimated kidney volumes appeared to be well correlated, but other groups have not yet validated these methods. In addition, the midslice method was developed in a cohort that included only patients with creatinine clearance $>$ 70 mL/min. In general, such patients have relatively small kidneys, making manual tracing measurement of TKV relatively easy, which may have influenced the results that were obtained. This method should therefore also be validated in patients with lower kidney function. Estimation methods to assess TKV may also be used in clinical trials, but only when they can accurately and reliably detect changes in TKV over time. To our knowledge, these issues have not been investigated to date.

Given these considerations, the objective of the present study was to investigate cross-sectionally these methods to estimate TKV in a patient group with a wide range of kidney function. Furthermore, we investigated in a longitudinal study whether these estimation methods can accurately detect changes in TKV.

METHODS

Patients and Study Design

For this study, all MR images of patients with ADPKD that were available from 2007 through 2014 were used. These patients participated in 1 of 3 studies that were performed by the departments of nephrology at the University Medical Centers of Groningen, Leiden, Nijmegen, and Rotterdam (all in the Netherlands). Details of the study protocols have been published elsewhere[4,8,9]; see Figure S1 (available as online supplementary material) for a flow diagram showing the assembly of the cohort. All patients were included if an MR image was available. ADPKD was diagnosed based on the modified Ravine criteria.[10] The Medical Ethics Committee of the University Medical Center Groningen approved the protocols of the 3 studies that were conducted in accordance with the International Conference of Harmonization Good Clinical Practice Guidelines and in adherence to the ethics principles that have their origin in the Declaration of Helsinki. All patients gave written informed consent.

Measurement and Collections

All participants collected a 24-hour urine sample the day preceding the MR imaging (MRI), in which urinary albumin concentration was measured. At the outpatient clinic on the day of MRI, blood pressure was assessed at rest in a supine position with an automatic device (Dinamap; GE Medical Systems) for 15 minutes and weight and height were determined. Blood samples were drawn for determination of creatinine level with an enzymatic assay (isotope-dilution mass spectrometry traceable; Modular; Roche Diagnostics), which was used to estimate glomerular filtration rate (GFR) using the CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) equation.[11,12]

MR Imaging

All participants underwent a standardized abdominal MRI protocol without the use of intravenous contrast. For the specific MRI protocol, see Item S1.

Gold-Standard Method: mTKV

Kidney and liver volumes were measured on the coronal fat saturated T2-single shot fast spin-echo sequence if possible. If the T2-weighted images showed too low quality, the MR image was excluded. Kidney and liver volumes were measured using the manual tracing method. Kidney and liver boundaries were manually traced using the commercially available software Analyze Direct 11.0 (Analyze Direct Inc). Kidney and liver volumes were calculated from the set of contiguous images by summing the products of the area measurements within the kidney or liver boundaries and slice thickness.[6] Nonrenal parenchyma (eg, the renal hilus) was excluded from measurement.

Estimation Methods: Estimated TKV

The 2 formulas used to estimate kidney volume were derived from the literature.[2,7]

We first used the midslice method to estimate TKV ($eTKV_{PANK}$).[7] The midslices of the coronal MR images were selected for each kidney separately. The midslice was defined as the slice for which the slice number corresponds to half the sum of the numbers of the first and last slice that contained the kidney. If the sum was odd, the midslice number was rounded up. $eTKV_{PANK}$ was calculated in milliliters, with midslice area and slice thickness in millimeters squared and millimeters, respectively. $eTKV_{PANK}$ was calculated as the sum of the left eKV_{PANK} (ie, 0.624 $\times$ midslice area $\times$ number of slices covering the left kidney $\times$ slice thickness/1,000) and right eKV_{PANK} (ie, 0.637 $\times$ midslice area $\times$ number of slices covering the right kidney $\times$ slice thickness/1,000).

Second, we used the ellipsoid method to estimate TKV ($eTKV_{ellipsoid}$).[2] For each kidney, length was measured as the average maximal longitudinal diameter measured in the coronal and sagittal plane. Width was obtained from the transversal image at maximum transversal diameter, and depth was measured from the same image perpendicular to the width measurement. $eTKV_{ellipsoid}$ was calculated in milliliters, with length, width, and depth all in millimeters. $eTKV_{ellipsoid}$ was calculated as the sum of the left $KV_{ellipsoid}$ and right $KV_{ellipsoid}$, both derived by the equation $\pi/6 \times (length_{coronal} + length_{sagittal})/2 \times width \times depth/1{,}000$. Of note, to assess $eTKV_{ellipsoid}$, no specific software is necessary, in contrast to assessment of mTKV and $eTKV_{PANK}$.

Table 1. Participants' Characteristics

	Whole Study Group (N = 220)	Patients With Follow-up (n = 48)	Test Set (n = 10)
Age, y	47.0 ± 8.6	39.2 ± 7.4	44.3 ± 10.2
Male sex	114 (51.8)	34 (71)	3 (30)
Body mass index, kg/m^2	26.9 ± 4.3	26.3 ± 3.4	27.1 ± 7.2
Body surface area, m^2	2.0 ± 0.2	2.1 ± 0.2	1.96 ± 0.2
Diastolic BP, mm Hg	82.2 ± 9.5	82.6 ± 8.8	85.4 ± 11.0
Systolic BP, mm Hg	132.7 ± 13.0	132.9 ± 11.6	134.1 ± 18.0
Antihypertensive medication	190 (86.4)	39 (81)	9 (90)
Plasma creatinine, mmol/L	125.5 ± 39.7	102.1 ± 31.7	127.4 ± 20.4
eGFR, mL/min/1.73 m^2	56.8 ± 20.3	79.7 ± 22.6	49.6 ± 10.2
24-h urine volume, L	2.36 ± 0.77	2.48 ± 0.87	2.60 ± 0.80
Albuminuria, mg/24 h	46.7 [21.2-88.2]	46.2 [19.0-181.0]	67.9 [17.0-95.4]
Kidney volume			
Total, L	1.96 [1.28-2.82]	1.79 [1.36-2.56]	1.78 [1.37-2.86]
Left, L	1.00 [0.67-1.52]	0.99 [0.73-1.39]	0.92 [0.70-1.62]
Right, L	0.92 [0.60-1.38]	0.80 [0.57-1.17]	0.91 [0.67-1.24]
Liver volume, L	2.74 [1.73-3.07]	NA	1.76 [1.62-3.64]

Note: Values for categorical variables are given as number (percentage); values for continuous variables, as mean ± standard deviation or median [interquartile range].

Abbreviations: BP, blood pressure; eGFR, estimated glomerular filtration rate; NA, not available.

Statistical Analyses

All analyses were performed with SPSS, version 22.0 (SPSS Inc). Normality of data was assessed by drawing Q-Q plots. Normally distributed variables are expressed as mean ± standard deviation, whereas non-normally distributed variables are given as median with interquartile range (IQR). Baseline characteristics of the study population are given overall (Table 1) and stratified for estimated GFR (eGFR) < 60 and ≥ 60 mL/min/1.73 m^2 (Table S1).

Differences between groups were tested using a 2-sample *t* test for normally distributed and Mann-Whitney *U* test for non-normally distributed data. For paired analyses, paired *t* test was used for normally distributed and Wilcoxon signed rank test was used for non-normally distributed data. McNemar test was used for paired nominal data. A 2-sided $P < 0.05$ was considered to indicate statistical significance.

In a test set of 10 patients stratified for kidney volume and MRI scanner, kidney volumes were measured and estimated twice by 4 reviewers (MDAvG, JvM, BvS, JvE). All reviewers were blinded to their previous results. Reproducibility was evaluated by assessing intra- and intercoefficient of variation (CV) for mTKV, $eTKV_{ellipsoid}$, and $eTKV_{PANK}$. The inter-CV was calculated for each of the 10 MR images as the standard deviation of TKV values assessed by all 4 assessors divided by the mean TKV of that image multiplied by 100%. The inter-CV given in this study is the mean of the inter-CVs of these 10 MR images. Intra-CV was calculated per MR image for each of the 4 assessors as the standard deviation of TKV values divided by the mean TKV multiplied by 100%. Per assessor, an average intra-CV was calculated. The intra-CV given in this study is the mean intra-CV (plus standard deviation) of these 4 assessors. We used paired *t* test to compare CVs between mTKV and eTKV.

To investigate whether eTKV correlated with mTKV, orthogonal regression analysis was performed, and Lins' concordance correlation coefficient was calculated using all MRI scans of our cohort.[13] Orthogonal regression uses the least square data modeling technique in which observational errors in both dependent and independent variables are taken into account. Agreement between eTKV and mTKV was evaluated by Bland-Altman analyses, with calculation of agreement limits (95% confidence interval). We used manual tracing as the gold standard for TKV measurement on the x-axis. Performance of the estimation methods compared with mTKV was assessed using bias, precision, and accuracy. For cross-sectional analyses, bias is expressed as mean percentage difference ([mTKV − eTKV]/mTKV × 100%), with positive values indicating underestimation of mTKV. Precision was defined as 1 standard deviation of bias. Accuracy was calculated as the percentage of eTKV values within 10%, 15%, and 20% of mTKV [P_{10}, P_{15}, and P_{20} respectively]). To investigate whether bias is dependent on patient or MR image characteristics, we performed regression analyses between bias and various variables; that is, age, length, body mass index, liver volume, and T1/T2-weighted images in univariate analyses. Differences in bias among the various scanners that were used were tested with analysis of variance. As standard quality control, ∼10% of all MRI scans were measured twice for mTKV, and this is referred to as $mTKV_{repeat}$. This was done to ensure that the observers maintained low interobserver variability. These scans were used to assess the precision and bias of mTKV.

To investigate whether the estimation methods can accurately detect changes in TKV, data for patients who had follow-up MR images available were used. For these longitudinal analyses, bias is expressed as the percent change in mTKV less the percent change in eTKV. Importantly, all follow-up scans were performed at the same MRI scanner as at baseline, and TKV was measured and estimated using the same series of images as at baseline, by reviewers blinded for baseline results.

To assess the consequences of using eTKV instead of mTKV, 2 analyses were performed. First, the effect on classification based on disease prognosis was assessed. To assess prognosis for clinical care, a classification system is used that categorizes patients into 5 classes based on thresholds for height-corrected TKV at a given age (A through E, with A indicating the best and E indicating the worst prognosis with respect to future kidney function decline).[2] In addition, there is a classification indicating whether a patient is suitable for inclusion in clinical trials. This classification contains 3 classes: patients who should not be included in clinical trials [I], patients whose suitability should be re-evaluated at yearly intervals [II], and patients who are optimal candidates for clinical trials [III]).[2] To assess reclassification, we created 5 × 5 and 3 × 3 cross-tabulations using height-corrected

TKV limits for their specific age.[2] In these tables, the proportion of reclassified participants was calculated when using height-corrected eTKV instead of height-corrected mTKV. For this analysis, only the "typical cases" were used, as advised for this classification system, defined as MR images with cysts with bilateral and diffuse distribution, in which all cysts contribute similarly to TKV.[2] Second, we assessed what the consequences were for sample size calculation for clinical trials using change in eTKV instead of change in mTKV. Sample size calculations were based on the literature[14] and used data from all patients who had longitudinal follow-up data available with respect to change in mTKV and eTKV. The number of patients needed per group was calculated assuming a power of 80% and 2-sided α of 0.05 to detect a percentage difference in TKV growth between treatment groups.[15]

RESULTS

Study Participants

The study population consisted of 220 patients with ADPKD; their characteristics are listed in Table 1. We excluded 44 patients because no T2-weighted images were available to perform both estimation methods. The patients were relatively young, with a mean age of 47.0 ± 8.6 (standard deviation) years, and already showed clear signs of disease. Most patients used antihypertensive medication. eGFRs were decreased (56.8 ± 20.3 [range, 17.0-129.2] mL/min/1.73 m^2). Urinary albumin excretion (46.7 [IQR, 21.2-88.2] mg/24 h) and TKV (1.96 [IQR, 1.28-2.82] L) were increased.

Reproducibility of mTKV and eTKV

Table 2 shows a test set for assessing reproducibility. Average intraobserver CVs were 1.8% for mTKV and 2.6% for total liver volume, whereas interobserver CVs were 2.3% and 3.5%, respectively. Variability for eTKV$_{ellipsoid}$ was significantly higher than for mTKV, whereas for eTKV$_{PANK}$, no significant differences were found when compared to mTKV. Analysis time was approximately 55 minutes per MR image for mTKV and 65 minutes for total liver volume, with higher analysis times in case of larger organs. Average time needed per MR image to estimate TKV using the midslice method was 15 minutes; using the ellipsoid method, 5 minutes.

Table 2. Test Set for Assessing Reproducibility

	Both Kidneys	Left Kidney	Right Kidney
mKV			
Intraobserver CV	1.8	2.3	1.9
Interobserver CV	2.3	2.6	2.9
eKV$_{ellipsoid}$			
Intraobserver CV	3.9[a]	4.9[a]	4.3[a]
Interobserver CV	6.3[a]	6.0[a]	8.5[a]
eKV$_{PANK}$			
Intraobserver CV	3.0	3.8	3.1
Interobserver CV	3.4	4.2	3.1

Note: Values are given as percentage. Intra- and interobserver CVs for mKV and for eKV$_{ellipsoid}$ and eKV$_{PANK}$. All CVs were calculated based on 10 patients.

Abbreviations: CV, coefficient of variation; eKV$_{ellipsoid}$, estimated kidney volume using ellipsoid method; eKV$_{PANK}$, estimated kidney volume using midslice method; mKV, measured kidney volume.

[a]$P < 0.05$ for difference in intra- or interobserver CV eKV versus corresponding value of mKV.

Performance of the TKV Estimation Methods

In the cohort for cross-sectional analyses, correlations of mTKV versus mTKV$_{repeat}$, eTKV$_{ellipsoid}$, and eTKV$_{PANK}$ are shown in Fig 1. Figures S2 and S3 show these correlations for left and right kidneys, separately. High correlations were observed for all 3 methods (mTKV$_{repeat}$: $r = 0.998$ [$P < 0.001$]; eTKV$_{ellipsoid}$: $r = 0.989$ [$P < 0.001$]; and eTKV$_{PANK}$: $r = 0.990$ [$P < 0.001$]. Figure 1 also shows Bland-Altman plots of mTKV versus the percentage difference between mTKV and mTKV$_{repeat}$ and both eTKV methods. mTKV$_{repeat}$ showed low bias (mean, 0.02% ± 3.2%). eTKV also did not systematically over- or underestimate mTKV (bias of 1.4% ± 9.2% and 4.6% ± 7.6% for eTKV$_{ellipsoid}$ and eTKV$_{PANK}$, respectively; Table 3). Bias for eTKV$_{PANK}$ was significantly higher than for mTKV$_{repeat}$ ($P = 0.005$), whereas bias for eTKV$_{ellipsoid}$ did not significantly differ from that for mTKV$_{repeat}$ ($P = 0.4$). Given the lower standard deviation, mTKV$_{repeat}$ had better precision and therefore better performance compared with eTKV$_{ellipsoid}$ and eTKV$_{PANK}$.

In addition, when these analyses were repeated with patients with ADPKD stratified for eGFR, we observed no significant difference in bias for eTKV$_{ellipsoid}$ and mTKV$_{repeat}$ in patients with eGFRs ≥ 60 mL/min/1.73 m^2 and eGFRs < 60 mL/min/1.73 m^2 ($P = 0.2$ and $P = 0.3$, respectively). Between eTKV$_{PANK}$ and mTKV$_{repeat}$, we also observed no significant difference in patients with eGFR < 60 mL/min/1.73 m^2 ($P = 0.2$) and those with eGFR ≥ 60 mL/min/1.73 m^2 ($P = 0.9$). Table S2 shows bias and accuracy for eTKV stratified by eGFR.

When investigating factors associated with bias, it appeared that liver volume was associated with bias in eTKV$_{PANK}$ ($P = 0.04$), but not with eTKV$_{ellipsoid}$ ($P = 0.1$). Bias was not associated with age ($P = 0.5$ and $P = 0.6$), height ($P = 0.8$ and $P = 0.1$), or strength of magnetic field ($P = 0.8$ and $P = 0.7$), respectively, for eTKV$_{ellipsoid}$ and eTKV$_{PANK}$.

Ability to Detect Changes in TKV When Using Estimation Methods

Follow-up data for TKV were available for 48 patients. Baseline characteristics for the longitudinal cohort are given in Table 1. These patients were

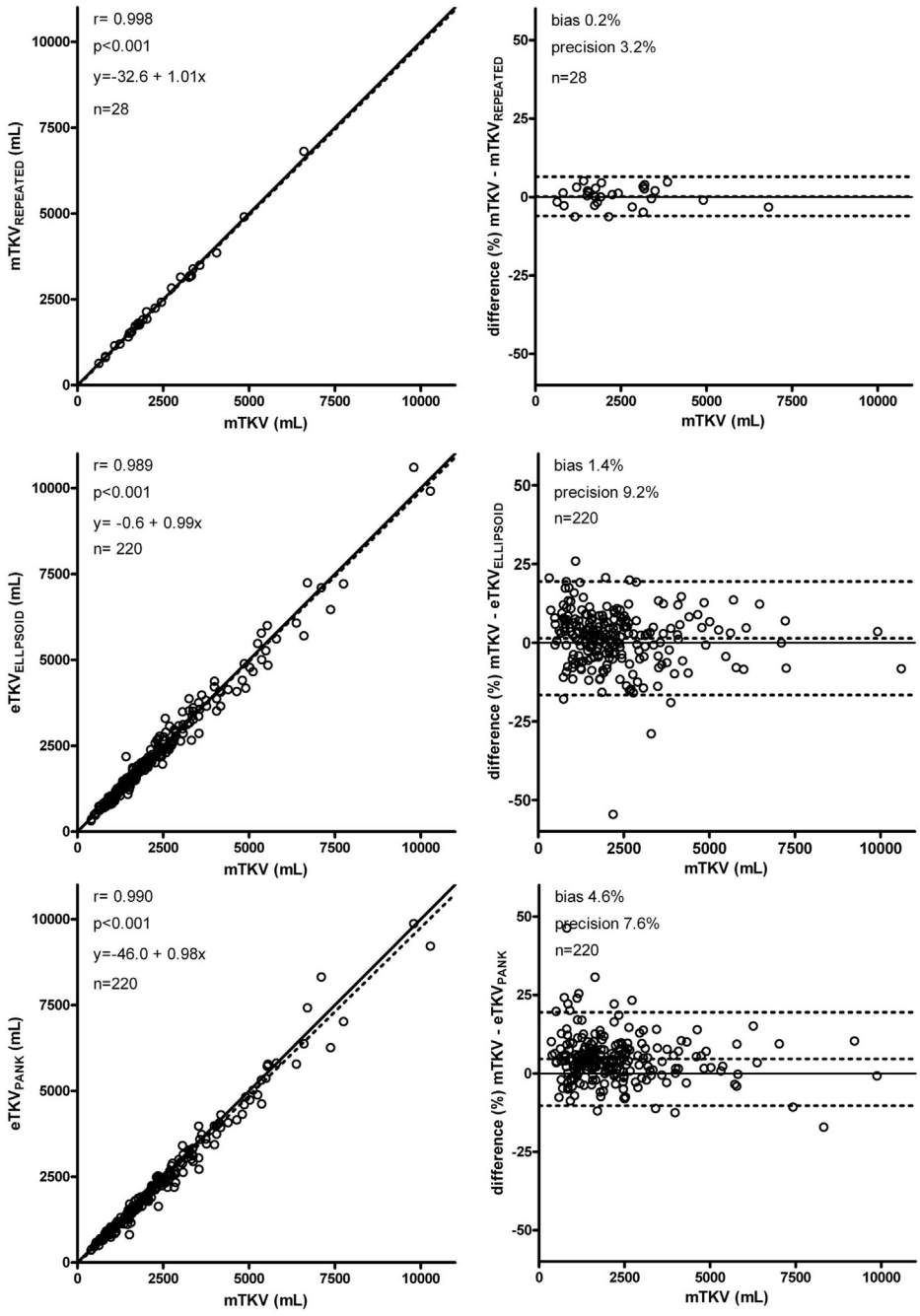

Figure 1. Cohort for cross-sectional analyses: associations between (upper panels) measured total kidney volume (mTKV) and repeat mTKV ($mTKV_{REPEATED}$) and estimated TKV using (middle panels) the ellipsoid method ($eTKV_{ELLIPSOID}$) and (lower panels) the midslice method ($eTKV_{PANK}$). (Left panel) scatter plots (solid line represents the line of identity, and the dotted line, actual regression line); (right panel) Bland-Altman plots (solid line indicates no difference, and dotted lines, mean difference [ie, bias] with 95% confidence interval).

younger, showed fewer signs of disease, and had higher eGFRs (79.7 ± 22.6 mL/min/1.73 m^2) but similar urinary albumin excretion (46.2 [IQR, 19.0-181.0] mg/24 h). During a follow-up of 3.0 years, mTKV increased from 1.79 (IQR, 1.36-2.56) to 2.18 (IQR, 1.55-2.73) L ($P < 0.001$). Median differences during follow-up were 0.25 (IQR, 0.04-0.54), 0.30 (IQR, 0.08-0.86), and 0.28 (IQR, 0.08-0.54) L for mTKV, $eTKV_{ellipsoid}$, and $eTKV_{PANK}$, respectively (Table 4). Change in eTKV compared to change in mTKV was not significantly different for both estimation methods ($P = 0.2$ and $P = 0.5$ for $eTKV_{ellipsoid}$ and eTKVPANK, respectively). Figure 2 plots percentage change in mTKV versus percentage change in eTKV. High concordance correlations were observed for $eTKV_{ellipsoid}$ ($r = 0.798$; $P < 0.001$) and

Table 3. Cohort for Cross-sectional Analyses: Performance of Ellipsoid and Midslice Methods for eKV

	$eKV_{ellipsoid}$ (n = 220)	eKV_{PANK} (n = 220)	mKV_{repeat} (n = 28)	*P* for mKV_{repeat} vs $eKV_{ellipsoid}$	*P* for mKV_{repeat} vs eKV_{PANK}
Left kidney volume, L	1.03 [0.65-1.48]	0.95 [0.63-1.45]	1.03 [0.75-1.78]	0.3	<0.001
Bias	−0.7	5.6	0.1	0.9	0.003
Precision	11.8	9.7	3.6		
Right kidney volume, L	0.90 [0.57-1.37]	0.88 [0.54-1.33]	0.98 [0.67-1.51]	0.003	<0.001
Bias	2.0	3.2	0.4	0.05	0.1
Precision	12.4	11.1	3.9		
Total kidney volume, L	1.93 [1.25-2.82]	1.81 [1.17-2.62]	1.92 [1.51-3.18]	0.004	<0.001
Bias	1.4	4.6	0.2	0.4	0.005
Precision	9.2	7.6	3.2		
Accuracy					
P_{10}	78.1	82.1	100	<0.001	<0.001
P_{15}	92.7	93.6	100	<0.001	<0.001
P_{20}	97.7	96.4	100	<0.001	<0.001
CCC	0.988	0.987	0.998		

Note: Values are given as percentage or median [interquartile range]. *P* values are calculated by paired *t* test when normally distributed, Wilcoxon signed rank test when non-normally distributed for continuous variables, and McNemar test for nominal variables.

Abbreviations and definitions: accuracy, percentage of eKV values within 10% (P_{10}), 15% (P_{15}), and 20% (P_{20}) of their corresponding mKV value; bias, mean percentage difference between mKV and eKV; CCC, concordance correlation coefficient; $eKV_{ellipsoid}$, estimated kidney volume using ellipsoid method; eTV_{PANK}, estimated kidney volume using midslice method; mTV_{repeat}, repeated measured kidney volume; precision, 1 standard deviation of bias.

$eTKV_{PANK}$ ($r = 0.866$; $P < 0.001$). Percentage change in eTKV did not show systematic under- or overestimation, with bias and precision (percent change mTKV − percent change eTKV) of −2.2% ± 10.3% and −1.8% ± 8.3% for $eTKV_{ellipsoid}$ and $eTKV_{PANK}$, respectively (Fig 2). In most patients, bias for change in eTKV was between −10% and 10% (72.3% and 74.5% of patients for $eTKV_{ellipsoid}$ and $eTKV_{PANK}$, respectively).

Consequences of Using eTKV Instead of mTKV

When using eTKV methods instead of mTKV for risk classification with respect to prognosis for rapid kidney function decline, we excluded the radiologically atypical ADPKD cases (n = 27), as advised for this classification system. There were 93.3% ($eTKV_{ellipsoid}$) and 90.2% ($eTKV_{PANK}$) of patients reclassified to their original risk categories (Table 5), whereas for both estimation methods, <1.6% of patients were reclassified to a higher risk category, and <8.5%, to a lower risk category. For classification for selection of patients for clinical trials, we observed that 97.4% ($eTKV_{ellipsoid}$) and 95.9% ($eTKV_{PANK}$) of patients were reclassified to their original categories. No patients were reclassified to a higher risk category when using $eTKV_{ellipsoid}$, and only 1 patient, when using $eTKV_{PANK}$ (Table 5).

The consequences of using percentage change in eTKV instead of percentage change in mTKV as the end point for sample size calculation for randomized controlled trials were assessed using data from the 48 patients with ADPKD for whom follow-up data were available. We calculated the number of study participants per treatment group needed to be enrolled to demonstrate a certain percentage decrease in rate of growth in TKV. Results are shown in Table S3. To detect, for instance, a 30% decrease in rate of growth in mTKV over 3 years, 186 patients are needed per treatment group, whereas for $eTKV_{ellipsoid}$ and $eTKV_{PANK}$ these numbers are 122 and 143, respectively.

DISCUSSION

This study was conducted to investigate whether TKV can be estimated accurately using the midslice (PANK) and ellipsoid methods in a group of patients with ADPKD with a wide range of kidney function. In a test set of 10 patients with ADPKD, we found that both estimation methods were highly reproducible. In our study cohort of 220 patients with ADPKD, both methods showed low bias, high precision, and high accuracy when compared to mTKV. This held for the overall cohort, as well as for patients with higher and lower eGFRs. In the 48 patients who had follow-up MR images available, change in eTKV was not different from change in mTKV for both methods.

Assessment of TKV using the gold-standard method of manual tracing is time consuming and needs specific software, which limits its applicability for clinical care. Methods have therefore been sought to estimate TKV in a more feasible way. Two

Table 4. Cohort for Longitudinal Analyses

	Baseline, L	Follow-up, L	Change, L	Change, %
Both kidneys				
mTKV	1.79 [1.36-2.56]	2.18 [1.55-2.73]	0.25 [0.04-0.54]	16.7 ± 17.1
$eTKV_{ellipsoid}$	1.86 [1.32-2.75]	2.39 [1.50-2.80]	0.30 [0.08-0.86]	19.3 ± 16.1
$eTKV_{PANK}$	1.79 [1.12-2.43]	2.03 [1.49-2.63]	0.28 [0.08-0.54]	17.8 ± 16.1
Left kidney				
mKV	0.99 [0.74-1.39]	1.23 [0.83-1.56]	0.13 [0.01-0.29]	15.0 ± 18.7
$eKV_{ellipsoid}$	1.03 [0.70-1.44]	1.26 [0.85-1.58]	0.10 [0.04-0.37]	17.7 ± 18.1
eKV_{PANK}	0.92 [0.68-1.24]	1.10 [0.78-1.44]	0.17 [0.04-0.36][a]	19.7 ± 19.0[a]
Right kidney				
mKV	0.80 [0.57-1.17]	0.99 [0.68-1.29]	0.13 [0.06-0.25]	19.4 ± 18.6
$eKV_{ellipsoid}$	0.81 [0.58-1.10]	1.04 [0.65-1.39]	0.14 [0.04-0.29]	23.1 ± 22.8
eKV_{PANK}	0.78 [0.60-1.14]	0.90 [0.65-1.24]	0.13 [0.04-0.24]	17.0 ± 19.6

Note: Baseline and follow-up (T)KV data for 48 patients with autosomal dominant polycystic kidney disease with follow-up data available. Values are given as mean ± standard deviation or median [interquartile range]. No significant differences between change in e(T)KV versus change in m(T)KV were noted, except for change in left eKV_{PANK} (as indicated with [a]).

Abbreviations: $e(T)KV_{ellipsoid}$, estimated (total) kidney volume using ellipsoid method; $e(T)KV_{PANK}$, estimated (total) kidney volume using mid-slice method; m(T)KV, measured (total) kidney volume.

[a] $P < 0.05$.

methods have been published recently[2,7]; however, they have not been validated to date. This formed the rationale to perform the present study. For determination of whether these estimation methods can be used to assess TKV, it is important to answer the following 5 questions.

First, it is important to investigate what the reliability of the gold-standard method is. In our study, we found that the variability in volumetric assessment by manual tracing was very low. In general, T1- instead of T2-weighted images are used for volumetry in ADPKD because researchers want to align with the original CRISP methodology. However, when the CRISP Study started, gadolinium-enhanced T1-weighted MR images were used. Because of the potential adverse effects of gadolinium, use of this contrast agent has since been discouraged. Bae et al[16] showed in 2009 that unenhanced T1-weighted volumes were significantly lower than contrast-enhanced T1-weighted volumes. These differences were more pronounced in smaller kidneys because in such cases, the ratio of kidney boundaries area to kidney volume is higher. Bae et al[16] mentioned that one should therefore contemplate using T2 MRI for quantification of TKV because the high kidney tissue contrast and hyperintense renal cysts in T2 images aid in delineating kidney boundaries against background tissues when compared to T1-weighted images. At that time, T2-weighted imaging required longer scanning time and was subjected to increased variation in image quality because of motion artefacts and was therefore not feasible. Nowadays, T2-weighted scanning time is shorter and respiratory triggering to avoid motion artefacts has become available. In our experience, this sequence has the best quality in visualizing polycystic kidneys. We therefore chose T2-weighted images (see Fig S4 for an example) instead of T1-weighted images for our study.

Second, do these estimation methods show low variability? Variability in mTKV versus $eTKV_{PANK}$ was not significantly different and satisfactorily low. Variability in $eTKV_{ellipsoid}$ was significantly higher compared to mTKV, meaning that this method is slightly more operator dependent than the midslice method, but still low. In line with this, reclassification to another risk category for rapid kidney function decline for clinical care (Irazabal classes A-E[2]) happened infrequently when using $eTKV_{PANK}$, as well as $eTKV_{ellipsoid}$ (Table 5). Given these results and because $eTKV_{ellipsoid}$ is more convenient (shorter duration per MR image and assessment possible using standard MRI software), we advise that $eTKV_{ellipsoid}$ be used rather than $eTKV_{PANK}$ for risk assessment in clinical care.

Third, does the estimation method show good agreement with the gold-standard method? We found for both estimation methods that eTKV correlated strongly with mTKV. Although bias and precision again showed better values for $mTKV_{repeat}$ (0.02% and 3.2%, respectively), results for $eTKV_{ellipsoid}$ and $eTKV_{PANK}$ were good. Bias was low for $eTKV_{ellipsoid}$ and $eTKV_{PANK}$ (1.4% and 4.6%, respectively), although for $eTKV_{PANK}$, it was slightly (but significantly) higher than for $mTKV_{repeat}$. In addition, precision was reasonable, now with slightly better results for $eTKV_{ellipsoid}$ ($eTKV_{ellipsoid}$ and $eTKV_{PANK}$: 9.2% and 7.6%, respectively; Table 3). Consequently, we found good accuracy for both estimation methods (P_{20} for $eTKV_{PANK}$ and $eTKV_{ellipsoid}$ of 96.4% and

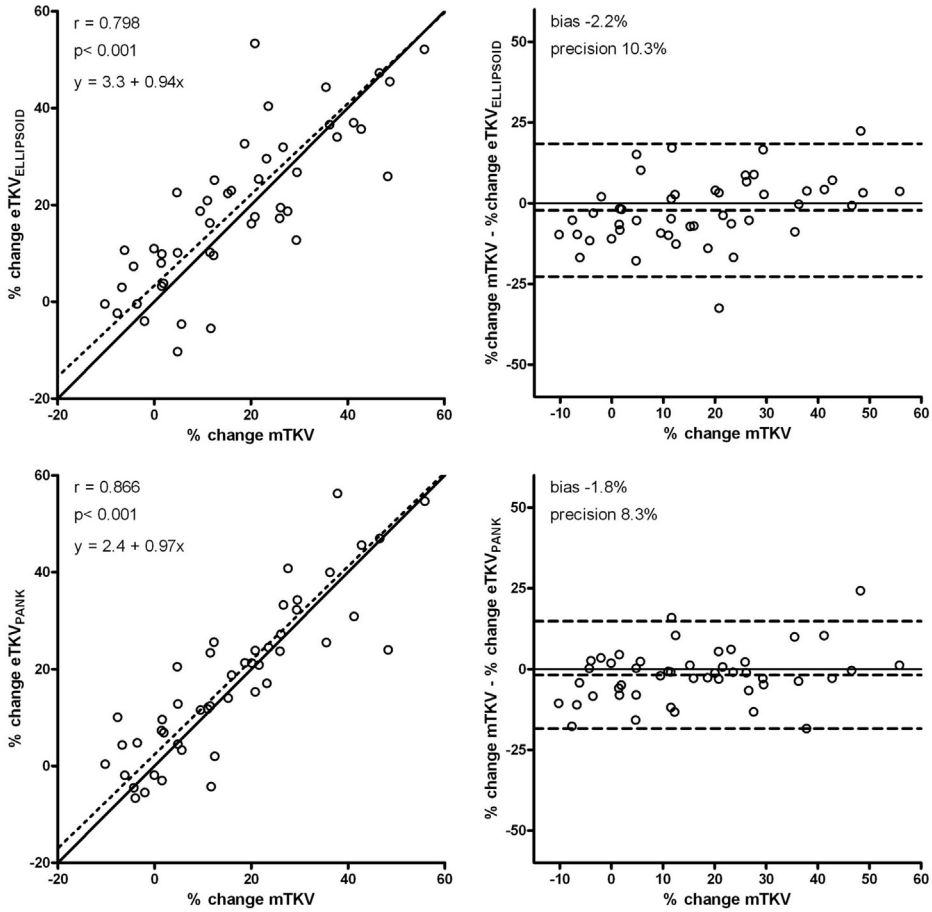

Figure 2. Cohort for longitudinal analyses: associations between percentage change in measured total kidney volume (mTKV) and percentage change in estimated total kidney volume (eTKV) using the ellipsoid method ($eTKV_{ELLIPSOID}$) and the midslice method ($eTKV_{PANK}$) in 48 patients with autosomal dominant polycystic kidney disease who had follow-up data available. (Left panel) scatter plots (solid line represents line of identity, and dotted line, actual regression line); (right panel) Bland-Altman plots (solid horizontal line indicates no difference, and dotted lines, mean difference [ie, bias] with 95% confidence interval).

97.7%, respectively). Our findings with respect to accuracy are consistent with values obtained in the cohort in which the ellipsoid method was developed (P_{10} of 70.3% vs 78.1% in the present study).[2] When stratified for kidney function, our results with respect to bias suggest that the midslice method may be less accurate in patients with ADPKD with lower kidney function, who generally have larger kidneys. Besides these statistical data, consequences for clinical care should be investigated when answering the question of whether estimation methods show good agreement with the gold-standard method. Irazabal et al[2] proposed a classification system for patients with ADPKD to assess their risk for rapid kidney function decline and to guide selection of patients for clinical trials. This classification system uses thresholds defined by age- and height-corrected TKV. We investigated the percentage of patients who are reclassified when using eTKV instead of mTKV. In the classification system for risk assessment, we observed that only a limited percentage of patients were reclassified, and these patients were most likely to be reclassified to a lower risk category (Table 5). No fundamental differences in results were observed for the 2 TKV estimation methods, and only one patient was reclassified when using $eTKV_{PANK}$ to a risk category that would preclude treatment (category B).

Fourth, can the estimation method detect changes in TKV over time? As far as we are aware, no study has yet investigated the performance of estimation methods to assess changes in TKV. In our analyses, we found a high concordance correlation between change in mTKV and change in $eTKV_{PANK}$ and $eTKV_{ellipsoid}$ during 3 years of follow-up, and no difference between change in mTKV and change in $eTKV_{PANK}$ and $eTKV_{ellipsoid}$ (Table 5). Consequently, when data for change in eTKV instead of change in mTKV are used, similar numbers of patients have to be included in clinical trials to be able to show a decrease in rate of growth in TKV (Table 5). These longitudinal results may seem surprising because they appear to be in contrast to our cross-sectional data, in

Table 5. Reclassification for Staging Into Risk Categories for Rapid Kidney Function Decline

Risk Category Classification

		$eTKV_{ellipsoid}$						$eTKV_{PANK}$				
		A	B	C	D	E		A	B	C	D	E
mTKV	A	5					A	4	1			
	B		28				B	1	27			
	C		5	66	2		C		6	65	2	
	D			4	47	1	D			6	45	
	E				1	35	E				3	33

Patient Selection for Trials

		$eTKV_{ellipsoid}$				$eTKV_{PANK}$		
		I	II	III		I	II	III
mTKV	I	5			I	4	1	
	II		28		II	1	27	
	III		5	155	III		6	150

Note: Based on Irazabal et al,[2] reclassification for staging into risk categories for rapid kidney function decline for clinical care (A-E) and for selection of patients for clinical trials based on thresholds for height-corrected TKV at a given age (I-III) using ellipsoid method ($eTKV_{ellipsoid}$) and using midslice method ($eTKV_{PANK}$) instead of mTKV.

Abbreviations: $eTKV_{ellipsoid}$, estimated total kidney volume using ellipsoid method; $eTKV_{PANK}$, estimated total kidney volume using midslice method; mTKV, measured total kidney volume.

which we showed that mTKV shows better reliability than $eTKV_{PANK}$ and $eTKV_{ellipsoid}$, albeit these differences were small. In our opinion, this may have 2 explanations. It could be that with eTKV methods, a systematic error is made in an individual patient in assessing TKV at baseline, for instance, due to a peculiar shape of a cystic kidney, but that the same error is made during follow-up because the shape of the cystic kidney has not changed. In this way, a systematic error in baseline eTKV will not translate in bias in change in eTKV during follow-up on a patient level. In addition, the natural variability in growth in TKV between patients may be so high that the limited variability that is added by using eTKV is not relevant when assessing mean change in TKV on a group level.

The fifth and last question to be answered is whether the estimation method is feasible from a clinical point of view. To estimate TKV using the midslice method, special software is necessary to measure the midslice area, limiting clinical applicability. In contrast, all clinicians can estimate TKV by the ellipsoid method using standard MR images without special software. Furthermore, the ellipsoid method requires less time to estimate TKV than using the midslice method, and both methods require far less time than assessment of mTKV with the gold-standard method of manual tracing.

The answers to these questions indicate that although eTKV may be slightly less precise than mTKV using the manual tracing method, it can be used with confidence in clinical care. Because numerically the 2 eTKV methods show hardly any differences with respect to bias, precision, and accuracy and no difference in ability to detect changes in eTKV, the more feasible ellipsoid method is to be preferred over the midslice method. Whether this conclusion is also valid for the use of $eTKV_{ellipsoid}$ instead of mTKV for clinical trials needs confirmation. To investigate this issue, results of these 2 assessment techniques should be compared in large-scale trials between different intervention groups using MR images obtained at baseline and during follow-up. Our data form the rationale to perform such studies.

A limitation of the present study is that our results hold primarily true for the cross-sectional correlation between mTKV and eTKV. Our results for follow-up data should be interpreted with caution because results are based on a limited number of patients. Strengths of this study are that we investigated both estimation methods in a group of patients with ADPKD with relatively well-preserved as well as reduced kidney function, and we are apparently the first to externally validate both estimation methods.

In conclusion, we demonstrated that both methods to estimate TKV perform relatively well in patients with ADPKD overall and in patients with preserved as well as reduced kidney function. In addition, both estimation methods detect relatively accurate changes in TKV over time. Because of these results and the higher feasibility of the ellipsoid method, we advise that the ellipsoid method be used for TKV estimation in clinical care. Whether this method can also be used for clinical trials deserves further study.

ACKNOWLEDGEMENTS

The DIPAK (Developing Intervention Strategies to Halt Progression of Autosomal Dominant Polycystic Kidney Disease) Consortium members are as follows (in alphabetical order throughout): Hedwig d'Agnolo, Niek F Casteleijn, Heleen Dekker, Joost Drenth, Johan W de Fijter, Ron T Gansevoort, Tom J Gevers, Hester Happé, Gert ter Horst, Peter Kappert, Esther Meijer, Dorien JM Peters (Consortium Leader), Remco Renken, H Pieterman, Mahdi Salih, Darius Soonawala, Edwin M Spithoven, Vicente E Torres, M Wasser, Jack FM Wetzels, and Robert Zietse. DIPAK 1 Study Steering Committee members: Joost Drenth, Johan W de Fijter, Ron T Gansevoort (Principal Investigator), Esther Meijer (Central Study Coordinator), Dorien JM Peters, Vicente E Torres, Jack FM Wetzels, and Robert Zietse. DIPAK 1 Data Safety Monitoring Board members: Carlo A Gaillard, Marjolein van Buren, Nick Veeger, and Marc Vervloet (Chair).

We acknowledge RL Kadijk for assistance at the outpatient clinic; L Schepel, J van Everdink, S Voorrips, C Plate, IL van Manen, MBR Wiertz, B van der Slik, and RR Buiten for measuring TKVs; and J Grozema and A Siebeijn-Kuiper for assistance during MRI; and thank the patients involved in this study for their participation.

Support: None.

Financial Disclosure: The authors declare that they have no relevant financial interests.

Contributions: Research area and study design: JPHD, JWdF, EM, DJMP, FWV, JFMW, RZ, RTG; data acquisition: EMS, MDAvG, ALM, NFC, PK, RJR; data analysis/interpretation: EMS, FWV, EM, CAG, RTG; statistical analysis: EMS, RTG; supervision or mentorship: CAG, RTG. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved. RTG takes responsibility that this study has been reported honestly, accurately, and transparently; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

SUPPLEMENTARY MATERIAL

Table S1: Baseline characteristics stratified by eGFR.

Table S2: Performance of $eTKV_{ellipsoid}$ and $eTKV_{PANK}$, stratified for eGFR < 60 and ≥ 60.

Table S3: Numbers of participants per group needed for RCTs to show a specific % difference in growth in 3-y TKV.

Figure S1: Flow diagram of the study design and classification.

Figure S2: Associations between mTKV and $mTKV_{repeat}$ and eTKV using $eTKV_{ellipsoid}$ and $eTKV_{PANK}$, left kidney.

Figure S3: Associations between mTKV and $mTKV_{repeat}$ and eTKV using $eTKV_{ellipsoid}$ and $eTKV_{PANK}$, right kidney.

Figure S4: Coronal image of a T2 single-shot, fast-spin, fat-saturated echo.

Item S1: Supplementary methods.

Note: The supplementary material accompanying this article (http://dx.doi.org/10.1053/j.ajkd.2015.06.017) is available at www.ajkd.org

REFERENCES

1. Grantham JJ, Torres VE, Chapman AB, et al. Volume progression in polycystic kidney disease. *N Engl J Med*. 2006;354(20):2122-2130.

2. 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.

3. ClinicalTrials.gov. Randomized Controlled Trial of Triptolide-Containing Formulation for Autosomal Dominant Polycystic Kidney Disease (ADPKD). https://clinicaltrials.gov/ct2/show/NCT02115659.

4. Meijer E, Drenth JP, d'Agnolo H, et al. Rationale and design of the DIPAK 1 study: a randomized controlled clinical trial assessing the efficacy of lanreotide to halt disease progression in autosomal dominant polycystic kidney disease. *Am J Kidney Dis*. 2014;63(3):446-455.

5. Chapman AB, Torres VE, Perrone RD, et al. The HALT polycystic kidney disease trials: design and implementation. *Clin J Am Soc Nephrol*. 2010;5(1):102-109.

6. Chapman AB, Guay-Woodford LM, Grantham JJ, et al. Renal structure in early autosomal-dominant polycystic kidney disease (ADPKD): the Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP) cohort. *Kidney Int*. 2003;64(3):1035-1045.

7. Bae KT, Tao C, Wang J, et al. Novel approach to estimate kidney and cyst volumes using mid-slice magnetic resonance images in polycystic kidney disease. *Am J Nephrol*. 2013;38(4):333-341.

8. Torres VE, Chapman AB, Devuyst O, et al. Tolvaptan in patients with autosomal dominant polycystic kidney disease. *N Engl J Med*. 2012;367(25):2407-2418.

9. Boertien WE, Meijer E, de Jong PE, et al. Short-term renal hemodynamic effects of tolvaptan in subjects with autosomal dominant polycystic kidney disease at various stages of chronic kidney disease. *Kidney Int*. 2013;84(6):1278-1286.

10. Pei Y, Obaji J, Dupuis A, et al. Unified criteria for ultrasonographic diagnosis of ADPKD. *J Am Soc Nephrol*. 2009;20(1):205-212.

11. Levey AS, Greene T, Kusek J, Beck GJ. A simplified equation to predict glomerular filtration rate for serum creatinine. In: ASN Program and Abstracts, 33rd Annual Meeting and 2000 Renal Week; October 10-16, 2000; Toronto, Ontario, Canada. Abstract 115A.

12. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150(9):604-612.

13. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45(1):255-268.

14. Noordzij M, Dekker FW, Zoccali C, Jager KJ. Sample size calculations. *Nephron Clin Pract*. 2011;118(4):c319-c323.

15. No author listed. Sample size calculator. http://www.stat.ubc.ca/~rollin/stats/ssize/. Accessed August 14, 2014.

16. Bae KT, Tao C, Zhu F, et al. MRI-based kidney volume measurements in ADPKD: reliability and effect of gadolinium enhancement. *Clin J Am Soc Nephrol*. 2009;4(4):719-725.