

# Limited Performance of Estimated Total Kidney Volume for Follow-up of ADPKD



Nathalie Demoulin<sup>1,2,7</sup>, Victoria Nicola<sup>3,7</sup>, Nicolas Michoux<sup>3</sup>, Valentine Gillion<sup>1,2</sup>, Thien Anh Ho<sup>4,5</sup>, Caroline Clerckx<sup>1,6</sup>, Yves Pirson<sup>1</sup> and Laurence Annet<sup>2,3</sup>

<sup>1</sup>Division of Nephrology, Cliniques universitaires Saint-Luc, Brussels, Belgium; <sup>2</sup>Institut de Recherche Expérimentale et Clinique, UCLouvain, Brussels, Belgium; <sup>3</sup>Department of Radiology, Cliniques universitaires Saint-Luc, Brussels, Belgium; <sup>4</sup>Division of Nephrology, CHU de Charleroi, Charleroi, Belgium; <sup>5</sup>Division of Nephrology, CHU de Tivoli, La Louvière, Belgium; and <sup>6</sup>Division of Nephrology, Clinique Saint-Pierre, Ottignies, Belgium

**Introduction:** Total kidney volume (TKV) is a qualified biomarker for disease progression in autosomal dominant polycystic kidney disease (ADPKD). Recent studies suggest that TKV estimated using ellipsoid formula correlates well with TKV measured by manual planimetry (gold standard). We investigated whether the ellipsoid formula could replace manual planimetry for follow-up of ADPKD patients.

**Methods:** Abdominal magnetic resonance images of patients with ADPKD performed between January 1, 2013, and June 31, 2019, in Saint-Luc Hospital, Brussels, were used. Two radiologists independently performed manual TKV (mTKV) measures and kidney axial measures necessary for estimating TKV (eTKV) using ellipsoid equation. Repeatability and reproducibility of axial measures, mTKV and eTKV, and agreement between mTKV and eTKV were assessed (Bland-Altman). Intraclass correlation coefficient (ICC) was used to assess agreement on Mayo Clinic Imaging Classification (MCIC) scores.

**Results:** 140 patients were included with mean age  $45 \pm 13$  years, estimated glomerular filtration rate (eGFR)  $71 \pm 31$  ml/min per  $1.73 \text{ m}^2$ , and mTKV  $1697 \pm 1538$  ml. Repeatability and reproducibility were superior for mTKV versus eTKV (repeatability coefficient 2.4% vs. 14% in senior reader, and reproducibility coefficient 6.7% vs. 15%). Inter-technique reproducibility coefficient (95% confidence interval [CI]) was 19% (17%, 21%) in senior reader. Inter-technique agreement on derived MCIC scores was very good (ICC = 0.924 [0.884, 0.949]).

**Conclusion:** TKV estimated using ellipsoid equation demonstrates poor repeatability and reproducibility compared with that of mTKV. Inter-technique agreement is also limited, even when measurements are performed by an experienced radiologist. Estimated TKV, however, accurately determines MCIC score.

*Kidney Int Rep* (2021) 6, 2821–2829; <https://doi.org/10.1016/j.ekir.2021.08.013>

KEYWORDS: ADPKD; ellipsoid equation; manual planimetry; TKV

© 2021 International Society of Nephrology. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

ADPKD is the most common inherited nephropathy and the fifth cause of kidney failure.<sup>1</sup> ADPKD is characterized by the progressive development of numerous cysts leading to kidney enlargement and impairment of kidney function. TKV is an early predictor of chronic kidney disease progression, unlike decline in glomerular filtration rate, which generally occurs late in ADPKD.<sup>2–6</sup> TKV has accordingly been approved by the US Food and Drug Administration and the European Medicines Agency as a qualified

biomarker for disease progression.<sup>7</sup> Measurements of TKV and change in TKV, respectively, are used for patient selection and evaluation of efficacy of pharmacologic treatments in ADPKD trials.<sup>8,9</sup> Also, the MCIC based on height-adjusted TKV at a given age is one of the criteria used in several countries for access to reimbursement of tolvaptan, a targeted therapy for rapidly progressive ADPKD.<sup>3,10</sup>

TKV is most often measured using magnetic resonance imaging (MRI). The gold standard technique for assessing TKV is the manual tracing technique (manual planimetry), whereby whole kidney contours are traced in contiguous kidney slices and surface areas calculated using a specialized software.<sup>3,9,11</sup> Stereology, which involves counting the number of intersections of a randomly positioned grid over the kidney, has been shown to be accurate when compared to manual

**Correspondence:** Nathalie Demoulin, Division of Nephrology, Cliniques universitaires Saint-Luc, Avenue Hippocrate 10, B-1200 Brussels, Belgium. E-mail: [nathalie.demoulin@uclouvain.be](mailto:nathalie.demoulin@uclouvain.be)  
<sup>7</sup>ND and VN contributed equally.

Received 4 March 2021; revised 14 July 2021; accepted 9 August 2021; published online 19 August 2021

planimetry.<sup>11</sup> However, both manual planimetry and stereology are very time consuming (up to 50 minutes per analysis) and, thus, only performed in certain centers and/or for research purposes and interventional studies.<sup>3,11</sup> Techniques to estimate TKV more rapidly have been developed and compared to manual planimetry and stereology. TKV estimated using the ellipsoid formula has been shown to correlate well with measured TKV, albeit with lower reproducibility, repeatability, and accuracy than measured TKV.<sup>3,8,9,12,13</sup> Longitudinal analysis of estimated versus measured TKV changes in subsets of the DIPAK and ALADIN cohorts suggested a relatively good correlation but insufficient precision to detect between-treatment changes in TKV.<sup>8,9</sup>

Measurement of TKV has become good clinical practice in the follow-up of ADPKD patients. We investigated whether the ellipsoid technique could replace manual planimetry in real-life follow-up of ADPKD patients, especially when performed by experienced radiologists. We analyze repeatability and reproducibility of TKV using manual planimetry and ellipsoid formula, and agreement between TKV using the 2 methods.

## METHODS

### Study Design

All consecutive abdominal MRI examinations of patients with ADPKD performed between January 1, 2013, and June 31, 2019, in Saint-Luc Hospital, Brussels, Belgium, were used for the study. ADPKD was based on the Pei/modified Ravine criteria and/or mutation screening.<sup>14</sup> MRI was performed as routine care for diagnostic and prognostic purposes. Images obtained from patients having had unilateral nephrectomy or having reached kidney failure (dialysis or transplantation) were excluded from the study (Figure 1). This was done because the volume of native polycystic kidneys decreases substantially following initiation of dialysis and transplantation.<sup>15–17</sup> The Saint-Luc Hospital's Ethical Committee approved the study (2019/03JUL/296).

### MRI

MRI scans were performed according to local practice. The following sequences were obtained for each patient: coronal T2 with 4-mm slices and axial T1 (DIXON 3 mm, reconstructed each 1.5 mm or LAVA 4.4 mm reconstructed each 2.2 mm). Two radiologists, a senior radiologist (LA: with 8 years of experience in TKV assessment) and a junior radiologist (VN: trained for the study), reviewed the MR images. They classified patients as having either typical (class 1) or atypical (class 2) images, based on previously established criteria<sup>3</sup>

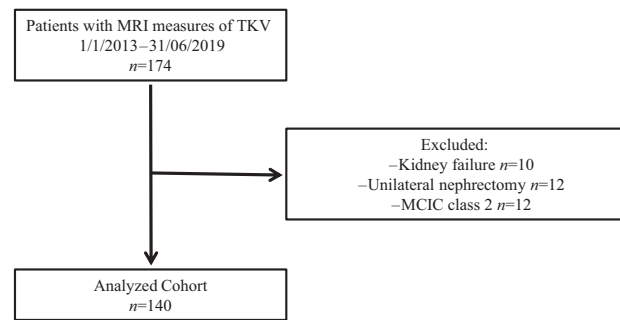


Figure 1. Study flow chart.

(Figure 1). Both radiologists reviewed cases for which their typical versus atypical classification differed to reach a consensus. Patients with typical images were then further classified into categories 1A to 1E according to height-adjusted TKV and age, by using the MCIC online calculator.<sup>3,18</sup>

### Measured and Estimated TKV

Measured TKV (mTKV) using the manual tracing method was performed using Vitrea (Toshiba) software. The kidney surfaces were manually drawn on the T1 axial images, and the volume was calculated by the software after extrapolation of the surfaces and manual correction.

Estimated TKV (eTKV) was obtained as follows: for each kidney, the length, width, and depth were measured on multiplanar reconstruction (MPR) obtained with T1 images. The length of the kidney was defined as the maximum length obtained on a coronal-oblique image after reconstruction following the sagittal plane of the kidney. The width was defined as the largest perpendicular axis obtained on the same coronal-oblique plane for the length, and the depth was the longest anteroposterior axis of the kidney perpendicular to the sagittal axis of the kidney. TKV was estimated using the ellipsoid technique, as follows:  $TKV = (\pi/6) \times L \times W \times D$  (L = maximum longitudinal length; W = maximum width perpendicular to L; D = maximum depth).<sup>12</sup>

The senior and junior radiologists (referred to as SR and JR hereafter) performed mTKV and eTKV measurements in 53 and 140 patients, respectively. They repeated mTKV and eTKV measurements in a subset of 10 and 22 patients stratified for kidney volume for repeatability analyses.

### Statistics

Bland-Altman methodology was used to assess (i) repeatability and reproducibility of axial measures, mTKV and eTKV, and (ii) intertechnique agreement between mTKV and eTKV. From each Bland-Altman plot were calculated the regression line of differences

(indicating proportional differences with the magnitude of the measurements), the mean bias (in %, indicating a potential under- or overestimation), and the limits of agreement (in %, indicating how well the measurements agree with each other). To ensure the validity of the definition of the limits of agreement (defined as  $\pm 1.96 \times$  standard deviation of paired differences around the mean bias), the normality of the paired differences was verified according to the Shapiro-Francia test ( $P < 0.05$ ). A 1-sample 2-sided  $t$  test was performed to assess whether the mean bias was significantly different from 0. A 2-sided  $t$  test based on the null hypothesis that the regression slope was equal to zero was performed.

The coefficient of variation was estimated as equal to standard deviation on the paired differences/mean (in %). The repeatability coefficient (assessing the intra-reader agreement), the reproducibility coefficient (assessing the interreader agreement), and the intertechnique reproducibility coefficient (assessing the intertechnique agreement) were derived from the limits of agreement (Supplementary Material, Methods section).<sup>19</sup> The upper limit of the 95% CI associated with the repeatability and reproducibility coefficients was then reported. This limit was used as the cutoff value above which a true variation in volume can be detected with a 95% confidence level, (i) by a given reader (if based on the upper limit of the 95% CI from the repeatability coefficient), (ii) regardless of the reader (if based on the upper limit of the 95% CI from the reproducibility coefficient), or (iii) regardless of the technique (if based on the upper limit of the 95% CI on the intertechnique reproducibility coefficient). The intertechnique reproducibility coefficient assessing the agreement between mTKV and eTKV was evaluated by JR ( $n = 53$ ) and SR ( $n = 140$ ).

Intertechnique agreement on the MAYO Imaging Classification (MCIC) score was evaluated with the ICC and its 95% CI. Strength of agreement was interpreted according to the Altman's scale as follows: ICC  $< 0.20$ , poor;  $0.21 < \text{ICC} < 0.40$ , fair;  $0.41 < \text{ICC} < 0.60$ , moderate;  $0.61 < \text{ICC} < 0.80$ , good; ICC  $> 0.81$ , very good.

A multivariable regression analysis was performed to identify potential independent prognostic factors associated with intertechnique agreement. A backward selection procedure was used to select the variables (eGFR, mTKV, hepatic volume, height, BMI). A variable was entered into the model if its associated  $P$  value was  $< 0.2$ , whereas it was removed from the model if its  $P$  value was  $> 0.4$ . The significance of the fit (given by the  $P$  value associated to the  $F$  test), the coefficient of determination adjusted for the number of independent variables entered into the model  $R^2_{\text{adjusted}}$ , and the regression equation were reported.

**Table 1.** Patient characteristics

Characteristic	Cohort ( $n=140$ )
Age, yr	45 $\pm$ 13
Male gender, n (%)	56 (40)
Body mass index	26 $\pm$ 5
eGFR, ml/min per 1.73 m <sup>2</sup>	71 $\pm$ 31
Right kidney volume, ml	803 $\pm$ 733
Left kidney volume, ml	894 $\pm$ 832
Total kidney volume, ml	1697 $\pm$ 1538
MCIC, 1A-1B-1B-1C-1D-1E	9-42-48-33-8
Total liver volume, ml	2316 $\pm$ 1209

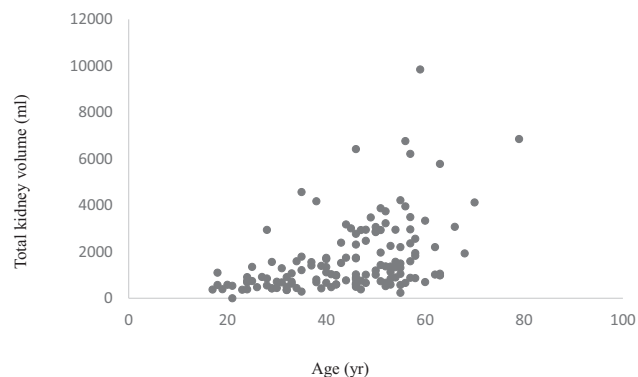
eGFR, estimated glomerular filtration rate; MCIC, Mayo Clinic Imaging Classification. Values are mean  $\pm$  SD unless otherwise noted.

The Medcalc software (Medcalc 19.0.3, Mariakerke, Belgium) was used to perform the analyses. Owing to the multiple comparisons that were performed, the significance level  $P$  of each statistical test was adjusted according to a Bonferroni correction.

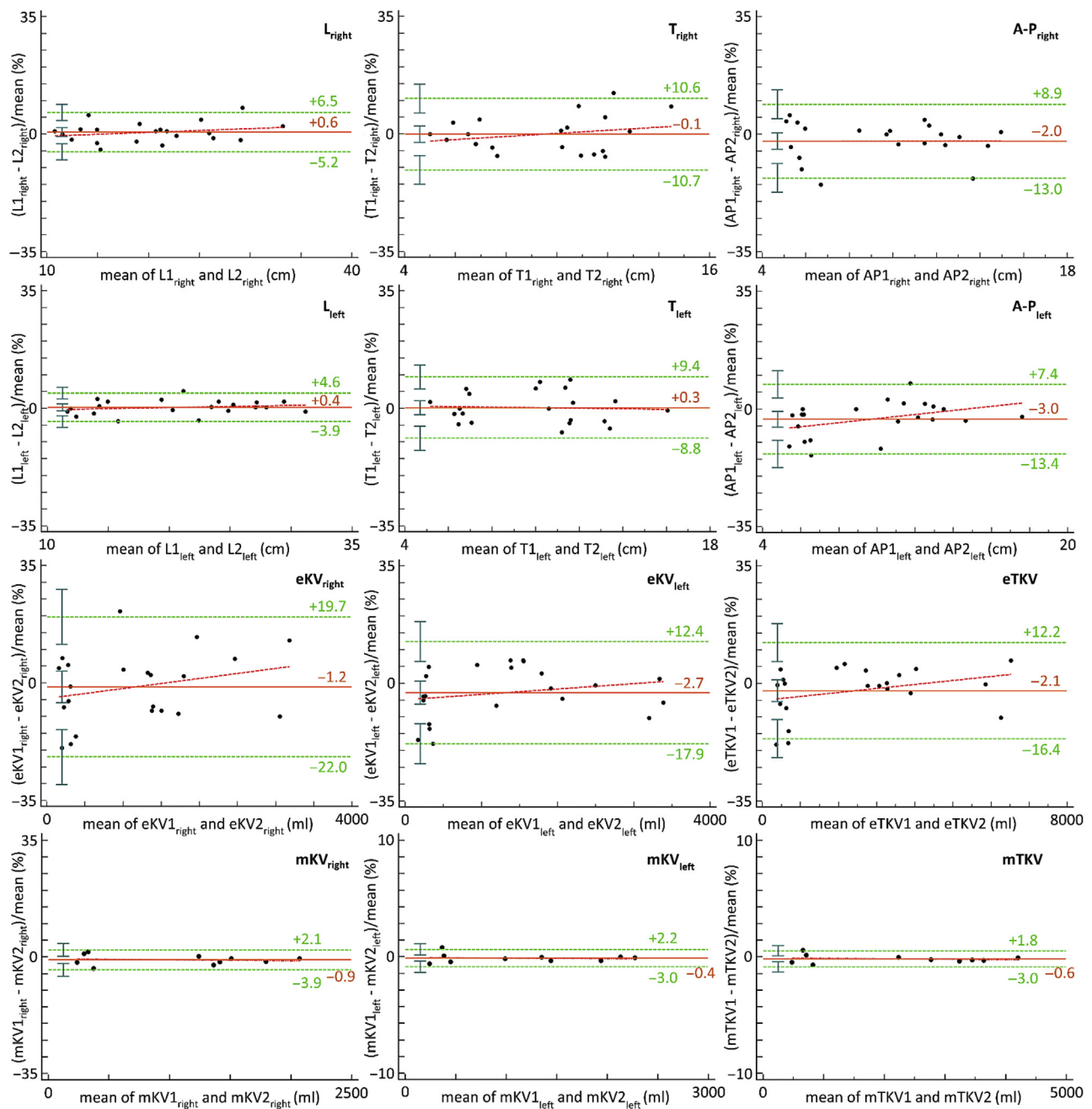
## RESULTS

### Patient Characteristics

One hundred forty patients were included in the study, after exclusion of 8 patients having reached kidney failure, 7 with unilateral nephrectomy, and 6 with atypical kidney imaging (MCIC class 2) (Figure 1). Mean age was 45  $\pm$  1 years and 40% were male. Patients had an eGFR (CKD-EPI [Chronic Kidney Disease Epidemiology Collaboration equation]) of 71  $\pm$  31 ml/min per 1.73 m<sup>2</sup> and mTKV of 1697  $\pm$  1538 ml (Table 1, Figure 2). Sixty-four percent of patients had MCIC of 1C-1E. Characteristics of patients for whom repeatability and reproducibility of mTKV and eTKV, and agreement between mTKV and TKV in JR, were analyzed ( $n=10$ ,  $n=22$ , and  $n=53$  respectively) were similar to that of the total cohort (Supplementary Table S1). Ten patients (7%) were treated with tolvaptan.



**Figure 2.** Total kidney volume distribution according to age in the total cohort ( $n = 140$ ).



**Figure 3.** Bland-Altman plots showing repeatability of axial measures, estimated TKV (eTKV) and measured TKV (mTKV) when performed by senior radiologist. Length (L), transverse diameter (T) (width), and anterior-posterior diameter (A-P) (depth) are expressed in millimeters, whereas eTKV is expressed in milliliters. The mean bias (orange solid line), regression line of differences (red dotted line), and limits of agreement (green dotted lines) with 95% CIs (green whiskers) are represented.

### Repeatability and Reproducibility of Axial Measures, eTKV and mTKV

A total of 36 Bland-Altman plots were used to assess the repeatability and reproducibility of axial measures, eTKV and mTKV (JR and SR). Figure 3 shows Bland-Altman plots for repeatability of kidney axial measures, eTKV ( $n = 22$ ) and mTKV ( $n = 10$ ), when performed by SR. Coefficient of variation, mean bias,  $P$  value associated with regression line of differences, limits of agreement, and repeatability and

reproducibility coefficients for the different analyses are reported in Tables 2 and 3.

Repeatability was superior for mTKV compared with eTKV (repeatability coefficient of 2.4% vs. 14% in SR and 4.6% vs. 17% in JR). Reproducibility was also superior for mTKV compared with eTKV (6.7% vs. 15%). All these differences were statistically significant at  $P < 0.05$  (nonoverlapping of the 95% CIs associated with the coefficients that are compared). JR and SR both spent approximately 50



**Table 2.** Repeatability of axial and kidney volume measures, using manual planimetry and ellipsoid formula

	L <sub>right</sub>	T <sub>right</sub>	A-P <sub>right</sub>	L <sub>left</sub>	T <sub>left</sub>	A-P <sub>left</sub>	eKV <sub>right</sub>	mKV <sub>right</sub>	eKV <sub>left</sub>	mKV <sub>left</sub>	eTKV	mTKV
Coefficient of variation (%)												
JR	1.9	3.7	8.7	2.8	4.9	6.4	13	3.8	10	2.2	8.7	2.3
SR	3.0	5.5	5.6	2.2	4.7	5.1	11	1.5	7.2	1.3	7.3	1.2
Bias <sup>a</sup> , %												
JR	+0.6	+1.9	+0.4	+0.7	-0.1	+0.3	+1.8	+1.2	+0.9	+2.3	+1.5	+1.7
SR	+0.6	-0.1	-2.0	+0.4	+0.3	-3.0	-1.2	-0.9	-2.7	-0.4	-2.1	-0.6
Regression <sup>b</sup>												
JR	0.008 <sup>c</sup>	0.050	0.278	0.063	0.305	0.390	0.032	0.793	0.165	0.282	0.013 <sup>c</sup>	0.843
SR	0.298	0.307	0.958	0.422	0.813	0.056	0.241	0.755	0.322	0.781	0.137	0.651
Limits of agreement, %												
JR	(-3.0, +4.2)	(-5.4, +9.2)	(-17, +17)	(-4.8, +6.2)	(-9.8, +9.6)	(-12, +13)	(-23, +27)	(-6.3, +8.6)	(-19, +20)	(-2.1, +6.7)	(-16, +19)	(-2.8, +6.3)
SR	(-5.2, +6.5)	(-11, +11)	(-13, +8.9)	(-3.9, +4.6)	(-8.8, +9.4)	(-13, +7.4)	(-22, +20)	(-3.9, +2.1)	(-18, +12)	(-3.0, +2.2)	(-16, +12)	(-3.0, +1.8)
Repeatability coefficient, %												
JR	3.6 (2.5, 4.7)	7.3 (5.0, 9.6)	17 (12, 22)	5.5 (3.9, 7.2)	9.7 (6.8, 13)	13 (8.7, 16)	25 (17, 32)	7.4 (4.1, 11)	20 (14, 25)	4.2 (2.4, 6.2)	17 (12, 22)	4.6 (2.5, 6.6)
SR	5.9 (4.1, 7.7)	11 (7.4, 14)	11 (7.7, 14)	4.3 (3.0, 5.6)	9.3 (6.5, 12)	10 (7.0, 13)	21 (14, 27)	3.0 (1.7, 4.3)	14 (10, 18)	2.6 (1.5, 3.8)	14 (10, 19)	2.4 (1.3, 3.5)

A-P, kidney anterior-posterior diameter (depth); eKV, estimated kidney volume; eTKV, estimated total kidney volume; JR, junior radiologist; L, kidney length; mKV, measured kidney volume; mTKV, measured total kidney volume; SR, senior radiologist; T, kidney transverse diameter (width).

<sup>a</sup>Bias: mean bias.

<sup>b</sup>Regression: *P* value associated with the regression line of differences.

<sup>c</sup>The regression slope is significantly different from 0.

Significance level associated with statistical tests: L, T, A-P, eTKV, and mTKV (*P* = 0.0083), eTKV<sub>total</sub> and mTKV<sub>total</sub> (*P* = 0.0167). Limits of agreement and repeatability coefficients are reported, with their 95% CIs in parentheses. No significant mean bias was found for any of the analyses.

and 15 minutes per mTKV and eTKV measurement, respectively.

### Agreement Between mTKV and eTKV

The intertechnique reproducibility coefficient assessing the agreement between mTKV and eTKV was 20% (95% CI 16%, 24%) and 19% (95% CI 17%, 21%) when performed by JR and SR respectively (Figure 4, Table 4). This means that a variation of eTKV can be considered as clinically significant only if it exceeds 21%, when performed by SR and considering mTKV as the gold standard. Proportional differences with the magnitude of TKV were observed for both readers (JR:

*P* = 0.0007, SR: *P* = 0.0002). Visual analysis of Figure 4 shows that for volumes <2000 ml, there is a trend to underestimate TKV using eTKV. This underestimation decreases when the volume increases.

Figure 5 shows, using a bar graph, the values of the upper limits of the 95% CI associated with repeatability and reproducibility coefficients of mTKV and eTKV, and with the intertechnique reproducibility coefficient. When performed by the senior radiologist (SR), a change in volume >3.5% in an individual patient can be detected with a 95% confidence using mTKV, whereas only a change >19% can be detected with the same confidence using eTKV. mTKV and eTKV

**Table 3.** Reproducibility of axial and kidney volume measures, using manual planimetry and ellipsoid formula

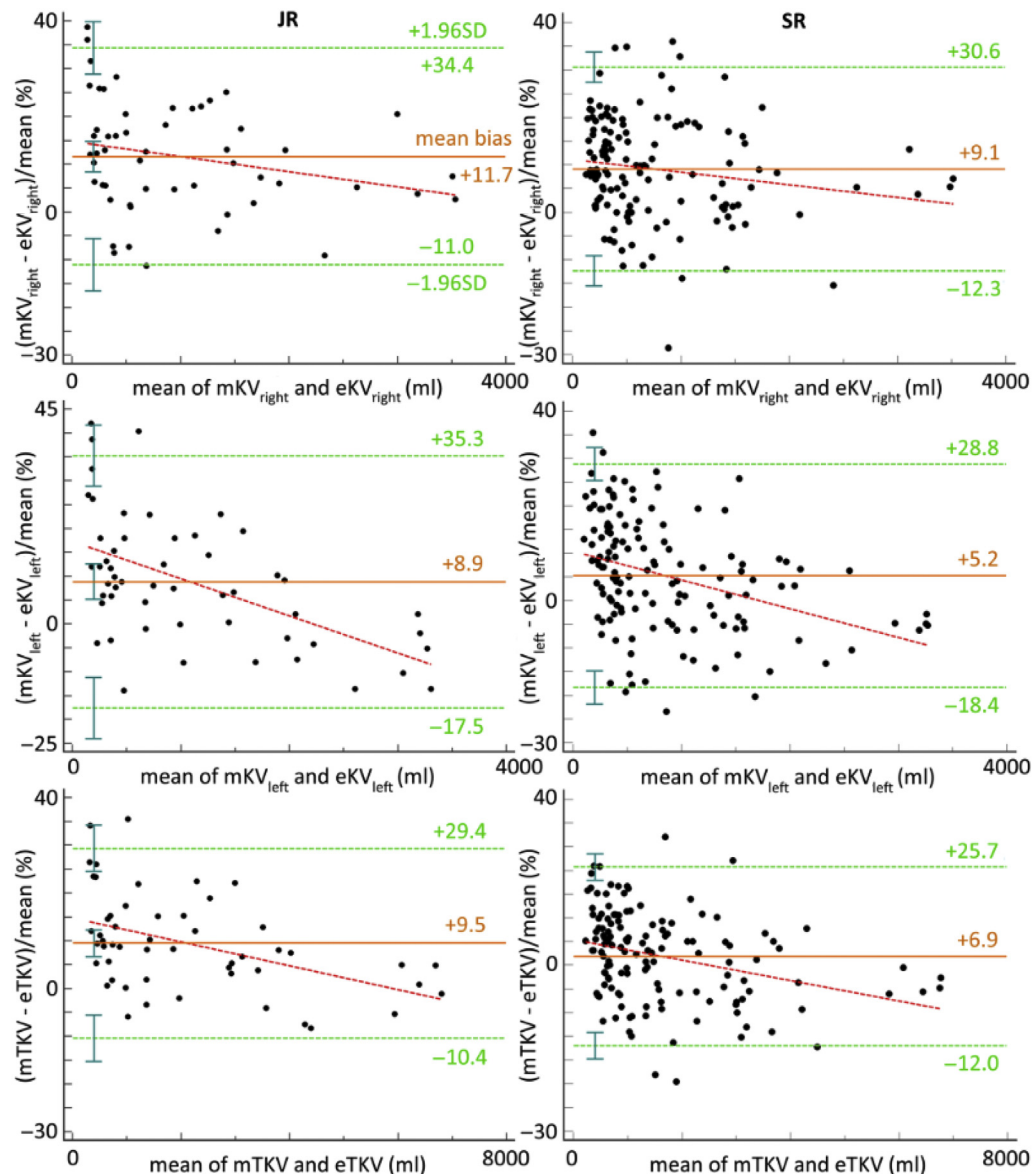
	L <sub>right</sub>	T <sub>right</sub>	A-P <sub>right</sub>	L <sub>left</sub>	T <sub>left</sub>	A-P <sub>left</sub>	eKV <sub>right</sub>	mKV <sub>right</sub>	eKV <sub>left</sub>	mKV <sub>left</sub>	eTKV	mTKV
Coefficient of variation, %												
JR	2.6	5.7	5.7	2.3	5.5	5.9	9.0	3.0	10	4.6	7.4	3.4
Bias <sup>a</sup> , %	+0.3	-0.3	+1.0	+0.8	-0.5	+0.8	+1.0	+1.0	+1.2	+3.2	+0.9	+2.2
Regression <sup>b</sup>	0.393	0.538	0.704	0.934	0.641	0.435	0.736	0.983	0.813	0.726	0.640	0.875
Limits of agreement, %												
JR	(-4.9, +5.5)	(-12, +11)	(-10, +12)	(-3.6, +5.3)	(-11, +10)	(-11, +12)	(-17, +19)	(-5.0, +6.9)	(-18, +21)	(-5.7, +12)	(-14, +16)	(-4.4, +8.9)
Reproducibility coefficient, %												
JR	5.2 (3.6, 6.7)	11 (7.9, 15)	11 (7.9, 15)	4.5 (3.1, 5.8)	11 (7.6, 14)	12 (8.0, 15)	18 (12, 23)	5.9 (3.2, 8.5)	20 (14, 25)	8.9 (4.9, 13)	15 (10, 19)	6.7 (3.7, 9.6)

A-P, kidney anterior-posterior diameter (depth); eKV, estimated kidney volume; eTKV, estimated total kidney volume; JR, junior radiologist; L, kidney length; mKV, measured kidney volume; mTKV, measured total kidney volume; SR, senior radiologist; T, kidney transverse diameter (width).

<sup>a</sup>Bias: mean bias.

<sup>b</sup>Regression: *P* value associated with the regression line of differences.

Significance level associated with statistical tests: L, T, A-P, eTKV, and mTKV (*P* = 0.0083), eTKV<sub>total</sub> and mTKV<sub>total</sub> (*P* = 0.0167). Limits of agreement and repeatability coefficients are reported, with their 95% CIs in parentheses. No significant mean bias was found for any of the analyses.



**Figure 4.** Bland-Altman plots for the agreement between estimated TKV (eTKV) and measured TKV (mTKV) for right kidney volume (right KV), left kidney volume (left KV), and TKV, when performed by junior and senior radiologists. Paired differences are expressed as a percentage of the geometric mean to overcome potential variability in the differences across the range of mean values. The mean bias (orange solid line), regression line of differences (red dotted line), and limits of agreement (green dotted lines) with 95% CIs (green whiskers) are represented.

can be interchanged with confidence when the variation in volume in a patient is  $>21\%$ .

### Agreement Between mTKV and eTKV on MCIC Score

Based on the results obtained by SR, agreement on MCIC scores derived from both techniques was very good (ICC = 0.924 [0.884, 0.949]). Twenty-one of 140 patients (15%) were misclassified by 1 risk class (Supplementary Table S2).

### Factors Contributing to the Bias Between Techniques

A linear regression model based on variables eGFR and BMI predicted the bias (mTKV - eTKV) / mean

(mTKV, eTKV) between both techniques. Results from the least squares multiple regression were as follows:  $P$  value<sub>F ratio</sub> = 0.0008,  $R^2$ <sub>adjusted</sub> = 0.0880,  $P$  value<sub>eGFR</sub> = 0.0197,  $P$  value<sub>BMI</sub> = 0.0049; regression equation: bias = 0.1439 + (0.0006 × eGFR) - (0.0046 × BMI); standard error<sub>eGFR</sub> = 0.0003, standard error<sub>BMI</sub> = 0.0016. A second model based on the variables mean TKV (computed from mTKV and eTKV) and BMI predicted the bias. Results from the regression were as follows:  $P$  value<sub>F ratio</sub> = 0.0001,  $R^2$ <sub>adjusted</sub> = 0.1221,  $P$  value<sub>mean TKV</sub> = 0.0009,  $P$  value<sub>BMI</sub> = 0.0168; regression equation: bias = 0.2003 - (0.00002 × mean TKV) - (0.0039 × BMI); standard error<sub>mean TKV</sub> = 0.000006, standard error<sub>BMI</sub> = 0.0016. Of note, a model based on the 3

**Table 4.** Intertechnique agreement (manual vs. ellipsoid formula) for kidney volume measures

	mKV <sub>right</sub> - eKV <sub>right</sub>	mKV <sub>left</sub> - eKV <sub>left</sub>	mTKV - eTKV
Coefficient of variation, %			
JR	12	13	10
SR	11	12	9.6
Bias <sup>a</sup> , %			
JR	+12 <sup>b</sup>	+8.9 <sup>b</sup>	+9.5 <sup>b</sup>
SR	+9.1 <sup>b</sup>	+5.2 <sup>b</sup>	+6.9 <sup>b</sup>
Regression <sup>c</sup>			
JR	0.118	<0.0001 <sup>d</sup>	0.0007 <sup>d</sup>
SR	0.0506	<0.0001 <sup>d</sup>	0.0002 <sup>d</sup>
Limits of agreement, %			
JR	(-11, +34)	(-18, +35)	(-10, +29)
SR	(-12, +31)	(-18, +29)	(-12, +26)
Intertechnique reproducibility agreement, %			
JR	23 (18, 27)	26 (21, 31)	20 (16, 24)
SR	22 (19, 24)	24 (21, 26)	19 (17, 21)

eKV, estimated kidney volume; eTKV, estimated total kidney volume; JR, junior radiologist; mKV, measured kidney volume; mTKV, measured total kidney volume; SR, senior radiologist.

<sup>a</sup>Bias: mean bias.

<sup>b</sup>Mean bias significantly different from 0.

<sup>c</sup>Regression: *P* value associated with the regression line of differences.

<sup>d</sup>Regression slope significantly different from 0.

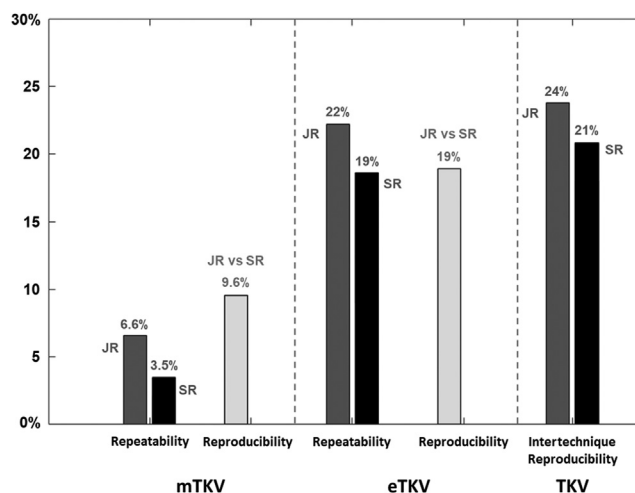
Significance level associated with statistical tests: mTKV and eTKV (*P* = 0.0125). Limits of agreement and agreement coefficients are reported, with their 95% CIs in parentheses.

variables, eGFR, mean TKV, and BMI, rejected the variable eGFR at *P* >0.4.

## DISCUSSION

Across a wide range of TKV, our study shows that TKV estimated using the ellipsoid equation has limited repeatability, reproducibility, and agreement with mTKV, even when performed by an experienced senior radiologist. Depending on the experience of the reader, the level of repeatability of eTKV ranged from 14% to 17%, well below that of mTKV ranging from 2.4% to 4.6%. Repeatability and reproducibility were lower for transverse and anterior-posterior axial measures of the kidneys, in comparison with length. Interestingly, kidney length has been shown to predict chronic kidney disease in ADPKD, with high levels of agreement between ultrasonographic and MRI scan kidney length.<sup>20</sup>

In terms of coefficient of variation (ie, of precision of measurements evaluated with a level of confidence of 68%), our results are consistent with those previously published.<sup>8,9,13</sup> However, assuming that a level of confidence of 95% is required in clinical practice, the upper limit of the 95% CI associated with repeatability and reproducibility coefficients should be used as a threshold value above which a clinically significant change in TKV in an individual patient can be assessed with confidence. As a result, when performed by the senior radiologist (SR), a change in volume >3.5% can



**Figure 5.** Upper limits of repeatability and reproducibility of measured TKV (mTKV) and estimated TKV (eTKV) and of agreement between mTKV and eTKV. Repeatability and agreement limits for junior radiologist (JR) are in light gray and for senior radiologist in dark gray; reproducibility limits are in black.

be detected with a 95% confidence using mTKV, whereas only a change > 19% can be detected with the same confidence using eTKV. These limits must be borne in mind as the TKV increases by an average of 5% to 6% per year throughout the course of ADPKD, irrespective of the causal mutation. TKV measurement using manual planimetry should, therefore, not be performed more frequently than every year for comparison purposes. The poor reproducibility of mTKV also limits the use of measurements performed by different radiologists, especially when using different MRI scans and software to calculate mTKV.

Agreement on kidney volume measures between mTKV and eTKV ranged between 17% and 21% (lower and upper limits of the 95% CI on intertechnique reproducibility coefficient in the senior reader). The underestimation of eTKV observed at smaller TKV was found to decrease when the mean TKV increases. Consequently, both techniques agree on the measure of change in TKV in an individual patient only when this change is superior to 21%. Again, considering the mean TKV annual growth rate of 5%, eTKV should not be used for comparative follow-up measures. Sharma *et al.* also showed that eTKV was not precise enough to be used for clinical studies to identify between treatment changes in TKV over a 1-year treatment period, albeit by substantially increasing the number of patients in the study.<sup>8</sup> Higashihara *et al.* have proposed the use of estimated height-adjusted TKV slope to estimate treatment effect over time.<sup>21</sup>

Estimated TKV did, however, accurately determine the Mayo Imaging Classification score, demonstrating a very good agreement with scores derived from mTKV. This is important as the MCIC is used worldwide to

evaluate the risk of progression and eligibility for approved treatments and interventional study protocols. The score can also be used to predict the decline in eGFR and renal survival in patients with typical ADPKD.<sup>3</sup>

Finally, the linear regression analysis showed that the variables mean TKV, eGFR, and BMI may influence the bias between mTKV and eTKV—the bias increasing when the eGFR increases and decreasing when the mean TKV or the BMI increases. However, the value of the coefficients of determination ( $R^2_{\text{adjusted}}$ ) of the models was well below 1.0, indicating that these regression models help little to explain the bias. Based on our results, we are unable to identify subgroups of ADPKD patients for whom the ellipsoid equation could reasonably estimate TKV.

The strengths of our study are the baseline TKV measures performed by both junior and senior radiologists and the real life setting. All our images were MRI as recommended in clinical practice for identification of ADPKD patients with progressive disease.<sup>10</sup> MRI is superior to computed tomography (CT) because of better cyst definition without requiring administration of contrast agents or radiation exposure.<sup>3</sup> T1- and T2-weighted images may be used for mTKV measures with similar repeatability and reproducibility, although T2-weighted images are more often of sufficient quality for TKV measures.<sup>22</sup>

The main limitation of this study was the inclusion of a limited number of patients with a potential influence on the value of the upper limits of the repeatability, reproducibility, and agreement coefficients. Most recently, semiautomated and automated measurements of kidney volumes from MRI scans have been developed using a deep learning–based approach and shown to be accurate in comparison with manual planimetry.<sup>4,23,24</sup> These techniques may have the potential to replace mTKV for baseline and follow-up measures of TKV in ADPKD patients, offering in theory a high repeatability and reproducibility of TKV measures. Other imaging techniques, such as those analyzing cystic load, vascularization, and texture of the kidney, may also refine measurement of ADPKD progression.<sup>25</sup>

Overall, we show that TKV estimated using the ellipsoid formula should not be used for follow-up of TKV in regard to low reproducibility and agreement with mTKV. Intertechnique agreement is limited even when measurements are performed by an experienced radiologist. Estimated TKV did, however, accurately determine the MCIC score. Semiautomated and automated measurements of kidney volumes from MR images may potentially replace mTKV in the near future in clinical routine.

## DISCLOSURE

All the authors declared no competing interests.

## ACKNOWLEDGMENTS

This study was supported by an unrestricted educational grant from Otsuka Pharmaceuticals (Nathalie Demoulin).

We wish to thank O. Devuyst for his constructive advice; V. Vanhole, C. Berghe, and N. Van Oost for their help with data collection; and S. Druart for the technical support.

## AUTHOR CONTRIBUTIONS

All authors contributed to the design of the study and review of the manuscript.

VN and LA reviewed all MR images and performed the kidney axial, eTKV, and mTKV measures.

ND wrote the manuscript.

NM performed statistical analysis.

## SUPPLEMENTARY MATERIAL

[Supplementary File \(PDF\)](#)

### Supplementary Methods.

**Table S1.** Patient characteristics according to analysis performed.

**Table S2.** Mayo Clinic Imaging Classification according to mTKV vs eTKV in the total cohort ( $n = 140$ ).

## REFERENCES

1. Chapman AB, Devuyst O, Eckardt KU, et al. Autosomal dominant polycystic kidney disease (ADPKD): executive summary from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. *Kidney Int.* 2015;88:17–27.
2. Grantham JJ, Torres VE, Chapman AB, et al. Volume progression in polycystic kidney disease. *N Engl J Med.* 2006;354:2122–2130.
3. 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;26:160–172.
4. Simms RJ, Doshi T, Metherall P, et al. A rapid high performance semi-automated tool to measure total kidney volume from MRI in autosomal dominant polycystic kidney disease. *Eur Radiol.* 2019;29:4188–4197.
5. Bae KT, Grantham JJ. Imaging for the prognosis of autosomal dominant polycystic kidney disease. *Nat Rev Nephrol.* 2010;6:96–106.
6. Chapman AB, Guay-Woodford LM, Grantham JJ, et al. Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease cohort: 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;64:1035–1045.
7. Perrone RD, Mouksassi MS, Romero K, et al. Total kidney volume is a prognostic biomarker of renal function decline and progression to end-stage renal disease in patients with



- autosomal dominant polycystic kidney disease. *Kidney Int Rep.* 2017;2:442–450.
8. Sharma K, Caroli A, Van Quach L, et al. Kidney volume measurement methods for clinical studies on autosomal dominant polycystic kidney disease. *PLoS One.* 2017;12, e0178488.
  9. Spithoven EM, van Castel MDA, Gansevoort RT, et al. Estimation of total kidney volume in autosomal dominant polycystic kidney disease. *Am J Kidney Dis.* 2015;66:792–801.
  10. 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;31:337–348.
  11. Magistrone R, Corsi C, Marti T, Torra R. A review of the imaging techniques for measuring kidney and cyst volume in establishing autosomal dominant polycystic kidney disease progression. *Am J Nephrol.* 2018;48:67–78.
  12. Higashihara E, Nutahara K, Okegawa T, et al. Kidney volume estimations with ellipsoid equations by magnetic resonance imaging in autosomal dominant polycystic kidney disease. *Nephron.* 2015;129:253–262.
  13. Shi B, Akbari P, Pourafkari M, Guiard E, Quist CF, Song X. Prognostic performance of kidney volume measurement for polycystic kidney disease. A comparative study of ellipsoid vs manual segmentation. *Sci Rep.* 2019;9:10996.
  14. Pei Y, Obaji J, Dupuis A, et al. Unified criteria for ultrasonographic diagnosis of ADPKD. *J Am Soc Nephrol.* 2009;20:205–212.
  15. Jung Y, Irazabal MV, Chebib FT, et al. Volume regression of native polycystic kidneys after renal transplantation. *Nephrol Dial Transplant.* 2016;31:73–79.
  16. Veroux M, Gozzo C, Corona D, et al. Change in kidney volume after kidney transplantation in patients with autosomal polycystic kidney disease. *PLoS One.* 2018;13:e0209332.
  17. Ishikawa I, Saito Y. Volume changes in autosomal dominant polycystic kidneys after the initiation of hemodialysis. *Nephron.* 1993;65:649–650.
  18. <https://www.mayo.edu/research/documents/pkd-center-adpkd-classification/doc-20094754>.
  19. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound Obstet Gynecol.* 2008;31:466–475.
  20. 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: CRISP Investigators. *Kidney Int.* 2015;88:146–151.
  21. Higashihara E, Fukuhara H, Ouyang J, et al. Estimation of changes in kidney volume growth rate in ADPKD. *Kidney Int Rep.* 2020;5:1459–1471.
  22. Van Gastel MD, Messchendorp AL, Kappert P, et al. T1 vs. T2 weighted magnetic resonance imaging to assess total kidney volume in patients with autosomal dominant polycystic kidney disease. *Abdom Radiol (NY).* 2018;43:1215–1222.
  23. Van Gastel MDA, Edwards ME, Torres VE, Erickson BJ, Gansevoort RT, Kline TL. Automatic measurement of kidney and liver volumes from MR images of patients affected by autosomal dominant polycystic kidney disease. *J Am Soc Nephrol.* 2019;30:1514–1522.
  24. Kline TL, Korfiatis P, Edwards ME, et al. Automatic total kidney volume measurement on follow-up magnetic resonance images to facilitate monitoring of autosomal dominant polycystic kidney disease progression. *Nephrol Dial Transplant.* 2015;31:241–248.
  25. Kline T, Korfiatis P, Edwards ME, Bae KT, Yu A, Chapman AB. Image texture features predict renal function decline in patients with autosomal polycystic kidney disease. *Kidney Int.* 2017;92:1206–1216.