Assessing Polycystic Kidney Disease in Rodents: Comparison of Robotic 3D Ultrasound and Magnetic Resonance Imaging

Nathan J. Beaumont,1 Heather L. Holmes,2 Adriana V. Gregory,3 Marie E. Edwards,3 Juan D. Rojas,1 Ryan C. Gessner,1 Paul A. Dayton,5 Timothy L. Kline,3,4 Michael F. Romero,2,4 Tomasz J. Czernuszewicz1,5

1SonoVol, Inc., Durham, NC, USA
2Department of Physiology & Biomedical Engineering, 3Radiology, 4Nephrology & Hypertension, Mayo Clinic, Rochester, MN, USA
5Joint Department of Biomedical Engineering, The University of North Carolina and North Carolina State University, Chapel Hill, NC, USA

Address for Correspondence

Address correspondence to Tomasz Czernuszewicz, Ph.D., SonoVol, Inc., 100 Capitola Dr. Suite 240., Durham, NC, 27713; Email: tomekc@sonovol.com; Phone: 844-766-6865 x703
Abstract

Polycystic kidney disease (PKD) is an inherited disorder characterized by renal cyst formation and enlargement of the kidney. PKD severity can be staged noninvasively by measuring total kidney volume (TKV), a promising biomarker that has recently received regulatory qualification. In preclinical mouse models, where the disease is studied and potential therapeutics are evaluated, the most popular noninvasive method of measuring TKV is magnetic resonance imaging (MRI). Although MRI provides excellent 3D resolution and contrast, these systems are expensive to operate, have long acquisition times, and consequently, are not heavily used in preclinical PKD research. In this study, a new imaging instrument, based on robotic ultrasound (US), was evaluated as a complementary approach for assessing PKD in rodent models. The objective was to determine the extent to which TKV measurements on the robotic US scanner correlated with both in vivo and ex vivo reference standards (MRI and Vernier calipers, respectively). A cross-sectional study design was implemented that included both PKD-affected mice and healthy wildtypes spanning sex and age for a wide range of kidney volumes. It was found that US-derived TKV measurements and kidney lengths were strongly associated with both in vivo MRI and ex vivo Vernier caliper measurements (R² = 0.94 and 0.90 respectively). In addition to measuring TKV, renal vascular density was assessed using acoustic angiography (AA), a novel contrast-enhanced US methodology. AA image intensity, indicative of volumetric vascularity, was seen to have a strong negative correlation with TKV (R² = 0.82), suggesting impaired renal vascular function in mice with larger kidneys. These studies demonstrate that robotic US can provide a rapid and accurate approach for noninvasively evaluating PKD in rodent models.
Introduction

Polycystic kidney disease (PKD) is an inherited disease defined by the development of many cysts in the kidneys. The disease is a genetic disorder with autosomal recessive (ARPKD) and autosomal dominant (ADPKD) forms. The autosomal dominant form is more common and has a prevalence of about 1 out of 2,500 people in developed countries. Cysts tend to increase in number and size over the lifetime of an individual, eventually causing chronic kidney disease (CKD) and possible kidney failure. To date, only one drug (Tolvaptan) is currently FDA approved to treat PKD and can only slow the progression of the disease. Thus, widespread research efforts to create better drugs for PKD are ongoing.

Rodent models play a crucial role in supporting disease research and drug testing environments. There are several different small animal models for PKD available today, and disease progression is typically measured terminally with histological methods. Histology provides cellular resolution and can quantify the presence of disease-relevant processes such as cystic diameter/distribution, inflammation, and fibrosis. However, histology is time consuming and ultimately destructive, limiting it to one time point per subject. In humans, estimated glomerular filtration rate (eGFR) is a common noninvasive biomarker used to assess renal function. However, in PKD, eGFR tends to decrease at late stages of the disease where cystic burden is substantial and difficult to treat. Due to the limitations of eGFR in patients, in 2015 the FDA qualified imaging-derived measures of total kidney volume (TKV) as a biomarker to quantify the efficacy of PKD drugs in clinical trials. In addition to earlier sensitivity for disease progression, TKV allows kidneys to be assessed throughout the onset and progression of PKD. In the context of preclinical research, TKV can provide improved statistics between groups with fewer animals required, compared to studies using invasive measurement approaches alone.
The most prevalent methods for noninvasive TKV measurement are magnetic resonance imaging (MRI) and ultrasound (US) because of their excellent soft tissue contrast, depth of penetration, and lack of ionizing radiation\textsuperscript{11}. Between the two, MRI has been considered the gold standard because of its high resolution, multiparametric pulse sequencing, and disease-relevant readouts for PKD progression such as TKV and cystic burden\textsuperscript{12–14}. However, MRI timepoints can be expensive due to high operational cost, long acquisition times, and limited access from high user demand. US imaging can provide a cost-effective alternative to MRI for rapid measurement of organ sizes, vascular density, and tissue stiffness, but has long suffered from issues of reproducibility due to its handheld form factor and user-dependence. Collecting high-quality reproducible rodent kidney measurements with US requires experienced sonographers and cumbersome workflows that can negate the cost, time, and throughput efficiencies that are promised by the modality.

Recently, a robotic preclinical US scanner has been developed to circumvent these challenges. This device uses an automated scanning mechanism to raster an US transducer across the entire body of a rodent, building up a widefield 3D image of the anatomy\textsuperscript{15}. By providing a single 3D image with a cohesive view of both kidneys within the context of the surrounding anatomy, we hypothesized that this technology could provide highly accurate TKV measurements in rodent models for disease. The following study tests this hypothesis in two cross-sectional studies with validation both \textit{in vivo} (i.e. MRI) and \textit{ex vivo} (i.e. Vernier caliper). Additionally, since PKD has been shown to decrease renal vascular density in rodent models\textsuperscript{16,17}, we evaluated whether this reduction in density could be assessed with a microbubble-contrast enhanced microvascular imaging mode provided by the robotic instrument (Acoustic Angiography)\textsuperscript{18}. Finally, intra and inter-rater reliability was assessed and reported.
Materials and Methods

**Imaging Studies**

Two different cohorts of mice were used in this study, one to compare the robotic US instrument with reference standard *in vivo* imaging systems, and one to compare the robotic US system to an *ex vivo* standard (Vernier calipers). Robotic US scans (Figure 1) were captured with a Vega imager (SonoVol, Inc., Durham, NC, USA), and compared with both MRI (16.4T Avance DRX 700WB, Bruker BioSpin, Billerica, MA, USA) and conventional preclinical US (Vevo 3100, FUJIFILM VisualSonics Inc., Toronto, Canada). All studies were approved by the University of North Carolina at Chapel Hill (UNC) and Mayo Clinic institutional animal care and use committees (IACUC). More details about scanning parameters can be found in the Online Supplementary Material and an example of robotic US scanning is provided in Supplementary Video 1.

Cohort 1: To evaluate the accuracy of the robotic US scanner against the reference standard *in vivo* imaging, a cross sectional study design was employed ensuring that kidneys over a range of sizes would be observed. Male and female PKD mice (N=7) were first imaged at the Mayo Clinic with both MRI and conventional US, per standard protocols, to serve as the gold standard. These animals were then transported to UNC for imaging with the robotic US instrument. Specific details on the imaging parameters for MRI and US can be found in the Supplemental Methods section. Additionally, physical and genetic characteristics of Cohort 1 are provided in Supplementary Table 1.

Cohort 2: To evaluate *in vivo* versus *ex vivo* size measurements, a second cohort (N=8) of healthy Nu/Nu mice were used (Charles River Labs, Wilmington, MA, USA). The cohort contained both female and male mice at two different ages (4 weeks old and 16 weeks old).
These animals were imaged exclusively with the robotic US scanner at UNC and were not transported to any other facility. Following robotic US acquisition of Cohort 2, all mice were euthanized to assess kidney size with Vernier calipers (Mitutoyo, Sakado, Japan) post necropsy. Additionally, the extracted kidneys were imaged on the robotic US scanner following Vernier measurements to confirm accuracy of in vivo measurements in the absence of surrounding anatomy. Excised kidneys were placed in warmed US gel on the Vega imager and scanned in 3D using the same imaging parameters as done in vivo.

**Image Analysis**

To determine kidney volume from a 3D image, kidneys were manually segmented from the data. Rapid segmentation from the robotic US data was performed using SonoEQ™ (SonoVol, Inc., Durham, NC, USA), which utilizes an open-source backend of 3D Slicer (www.slicer.org)¹⁹. To enhance kidney border delineation in the US data, the 3D volumes from both US transducers (dual element and linear array) were overlaid on one other with alpha blending. MRI and conventional US volumes were segmented by two separate readers using segmentation software developed in house²⁰. For the duration of image analysis, all readers were blinded to animal genotypes, sex, age, and one another’s kidney volume via mouse label randomization. For inter-modality assessment (US vs. MRI) and intra-modality (robotic US vs. conventional US) the segmentations from one expert reader were compared across all three categories: robotic US (R.C.G.), conventional US (H.L.H.), and MRI (M.E.E.). To determine accuracy of kidney sizing of the robotic US system compared to ex vivo measurements, all kidneys in Cohort 2 were assessed in two ways: with software calipers and 3D segmentations. Software calipers are linear measurement objects which can be placed within 3D images using SonoEQ software and were used to measure the length and width of the kidneys in the same orientation as the Vernier
calipers post necropsy. To assess inter-reader reliability for the robotic US scanner, four independent readers (N.J.B., J.D.R., R.C.G., T.J.C.) spanning a range of ultrasound imaging expertise segmented kidneys in the robotic US data and were compared. To assess intra-reader variability, one reader (N.J.B.) segmented the same kidney volumes again 67 days later without viewing previous segmentations. To quantify minimum detectable cyst size, a single coronal slice of kidneys with many cysts had digital caliper measurements made for 11 identifiable cysts.

**Assessing Relationship Between Kidney Vascularity and Kidney Size**

To assess the relationship between kidney size and kidney vascularity, an image intensity analysis was performed on Acoustic Angiography (AA) images of each kidney. The segmentations for each kidney from the B-mode images, described in the previous section, were applied to the co-registered AA volumes. This allowed the image voxels within the kidney to be analyzed as a histogram of intensity values. Because the brightness of a given pixel is correlated with the amount of microbubble contrast agents present in that localized region of tissue, the “% positivity”, or percentage of pixels within a region of interest above a given threshold is a proxy for vascularity. This metric of vascularity necessitated the selection of an intensity threshold above which pixels were identified as representing vasculature. By modulating the “vascularity threshold” across a range, the strength of the correlation between kidney size and vascularity could be assessed. This analysis was performed in Matlab R2017a (Mathworks, Natick, MA, USA).

**Statistical Analysis**

Agreement and correlation between imaging modalities was assessed in Matlab R2017a (Mathworks, Natick, MA, USA) using Ran Klein's Bland–Altman and Correlation Plot toolbox v1.10. Reported metrics included, Pearson correlation coefficient ($r$), Pearson squared ($r^2$),
coefficient of determination ($R^2$), Spearman correlation coefficient ($\rho$), line of best fit equation (least squares), coefficient of variation (CV), and limits of agreement (LOA). The linear regression equation was solved with a zero-intercept boundary condition. Bias between measurements was assessed using a Bland-Altman analysis. $P < 0.05$ was considered statistically significant. To quantify inter-reader reliability, the intra-class correlation coefficient (ICC) using the absolute agreement among measurements (“ICC(A,1)”)) formulation was computed across the four image readers as well as the CV between US measurements for each TKV.

Results

**Evaluating Kidney Size Measurement Accuracy (in vivo studies)**

Kidneys could be readily visualized in both MRI and robotic US datasets. As expected, the MRI datasets had better tissue contrast, but required more acquisition time per animal (5-10 min vs. 26 seconds). When compared side-by-side, the MRI and robotic US images had corresponding anatomical landmarks within each volume, such as cysts and blood vessels (Figure 2). Many cysts down to 0.4 mm in diameter were identifiable in US, but smaller cysts were not confidently distinguishable from US speckle (*Supplemental Figure S1*).

To assess the correlation between MRI and the robotic US scanner, segmented kidneys from Cohort 1 were compared via a linear regression and Bland-Altman analysis (Figure 3). The correlation between the segmentations between modalities was very strong, with an $R^2$ value of 0.94. Bland-Altman analysis revealed a mean underestimation bias of kidney volume by 4.9 mm$^3$ with robotic US, however it was not found to be significant ($p=0.62$). The LOA between the two modalities was 70 mm$^3$. Additionally, robotic US was compared to conventional US (*Supplemental Figure S2*), and conventional US to MRI (*Supplemental Figure S3*). When compared against conventional US, robotic US demonstrated excellent correlation ($R^2 = 0.97$)
and an even smaller LOA (44 mm³) as compared with MRI. Conventional US and MRI showed a high degree of correlation as well ($R^2 = 0.92$), albeit with a larger LOA (90 mm³).

**Evaluating Kidney Size Measurement Accuracy (ex vivo studies)**

To evaluate the relationship between *in vivo* kidney size measurements made on the robotic US scanner and *ex vivo* methods for sizing kidneys via Vernier calipers, a comparison study was performed on Cohort 2 including measurements for kidney width and length. For the robotic US scanner dataset, 3D images were acquired on both *in vivo* kidneys prior to necropsy (serving as the normal scenario for measurement in a longitudinal study), as well as explanted kidneys post necropsy (serving as the best case scenario for system performance and measurement accuracy). Figure 4 illustrates the placement of calipers (Vernier calipers in Panel A, as well as US calipers in B and C for *in vivo* and *ex vivo*, respectively). As measured by Vernier calipers, the range in kidney widths for this study was 5.42 to 8.03 mm, and the range in kidney lengths was 7.47 to 11.75 mm. Kidney dimensions measured with ultrasonic calipers, both *in vivo* and *ex vivo*, were strongly correlated with Vernier caliper measurements with coefficients of determinations $R^2 = 0.90$ and $R^2 = 0.95$, respectively (Figure 4). The expected trends for age and sex were also observed; older mice had larger kidneys than younger mice for the same sex, and males, on average, had larger kidneys than the females for a given age (Supplemental Figure S4).

**Evaluating Consistency Between Readers**

Figure 5 illustrates inter-reader variability in TKV for the Cohort 1 mice scanned on the robotic US system, sorted by size. The mean and standard deviation of the measurements are indicated, as well as the values of each of the four readers. The range in TKV for this cohort was 250.5 to 904.8 mm³. On average, the standard deviation for a TKV measurement across multiple readers was 43.62 mm³ with a minimum of 19.5 mm³ and maximum of 69.0 mm³ for animals 603 and
60A respectively. The mean coefficient of variation (stdev/mean) for these mice was 9.3%, with a min of 2.8% and max of 17.9% for mice 570B and 607, respectively. The ICC between the four readers was very strong at 0.93 (95% CI: 0.83 – 0.97). As for intra-user variability, the tested reader had an ICC of 0.96 (95% CI 0.88 – 0.99) between the two timepoints.

Evaluating the Relationship Between Kidney Vascularity and Kidney Size

It was hypothesized that in AA images there would be a correlation between the percentage of bright pixels within the kidney (% positivity, i.e. vascularity) and the kidney’s size. For instance, a kidney filled with large non-vascularized cysts would contain a high percentage of low intensity pixels within the kidney’s interior, and thus a low % positivity. Conversely, a kidney with high vascular density and a high % positivity. The optimal threshold yielding the maximum strength of correlation between vascularity and kidney size was an AA intensity value of 46 (a post-threshold image at this value can be seen in Figure 6A). At this threshold value, the R$^2$ value for the linear regression between vascularity and kidney size was 0.82. This analysis was exploratory in nature, and additional studies need to be performed to characterize the stability of this optimal threshold across different cohorts of animals, and how this can be used to extract other meaningful readouts from PKD models such as cystic burden.

Discussion

In this study, it was demonstrated that robotic 3D US can produce accurate and consistent in vivo measurements of TKV in rodent models. Correlation analyses between robotic US and MRI derived kidney volume yielded an R$^2$ value of 0.94 with no statistically significant bias. Furthermore, robotic US measurements were in very good agreement with conventional US measurements (R$^2 = 0.97$) and required substantially less time to acquire. This is a significant finding, as it illustrates that the new US instrument with a robotic form-factor can be used as a
high-throughput screening method for TKV, thus enabling large-cohort preclinical studies that would otherwise be cost/time prohibitive to perform with MRI or conventional US. When higher spatial resolution is desired, individual animals screened by US can always be selectively sent for MRI, similar to targeted recruitment in clinical studies. In addition to high correlation of kidney volume measurements between the modalities, cysts down to 0.4 mm in diameter were visible in both (Figure 2, Supplemental Figure S1). This suggests that longitudinal monitoring of total cyst volume, mean cyst size, and percent cyst space with robotic US could be possible. Other metrics to quantify PKD such as parenchymal fibrosis and renal blood flow (RBF) could also be measured with robotic US via shear wave elastography imaging (SWEI) and Doppler respectively. While these metrics were beyond the scope of our single time-point validation study, they will be explored in more detail in the future.

Additionally, it was demonstrated that wildtype kidney sizes measured *in vivo* with the robotic US system were highly correlated with kidney sizes measured *ex vivo* (Figure 4). *Ex vivo* organ sizing is a routine practice in preclinical research as noninvasive imaging is typically not conducted. Specifically, for the kidney, this can be done via volume displacement approaches or caliper measurements performed in either two axes (i.e. length and width) or three axes (i.e. length, width, height) that can be fit to a ellipsoidal equation to yield volume. In this study, *in vivo* measurements were compared against *ex vivo* measurements made by both ultrasonic and manual means (via Vernier calipers). In all cases, caliper measurements matched closely, with $R^2$ values ranging from 0.88 to 0.95.

Interestingly, a positive bias was observed when comparing *in vivo* US calipers to Vernier calipers (0.87 mm, $p<0.01$, Figure 4D). In order to determine if this was an artifact of the US reader seeing the border of the kidney incorrectly within the *in vivo* US data, we analyzed the
measurements between Vernier calipers and *ex vivo* US calipers where the tissue borders have maximum contrast against the coupling gel. In this case (Figure 4E), nearly the same overestimation of 0.88 mm by US can be seen (p<0.01). One explanation could be that the Vernier calipers were physically compressing the kidneys by this distance causing an underestimation in these data, which is a known challenge with using Vernier calipers for measuring tissue samples. When comparing the *in vivo* vs. *ex vivo* US calipers, measurements were highly consistent with no significant bias (Figure 4F).

This study found that kidney vascularity as measured by bright contrast-enhanced voxels, had a strong negative correlation to kidney volume as measured by US scans (R² = 0.82) (Figure 6). While microbubbles are generally considered safe for use both clinically and preclinically, using them in concert with high intensity ultrasound pulses can result in unwanted bioeffects and cellular damage. In rodent kidneys, the mechanical index (MI) threshold at which microbubble damage occurs has been found to depend on frame rate and imaging time of an area. One study found that imaging one location for 60 seconds at 1 frame per second with an MI above 0.6 can cause glomerular hemorrhage. Another larger study found no glomerular hemorrhage at an MI up to 1.9, but did find blood urea nitrogen increase with pulses at 1.0 MI. Because the Vega system performs acoustic angiography at an MI of 0.54, and rapidly translates the transducer across the kidney, it is not expected that significant bioeffects, as reported in previous studies, would occur from a kidney vasculature scan, however this hypothesis should be rigorously tested in future studies.

PKD in humans usually causes hypertension, and cardiovascular disease is the most common cause of death for PKD patients. In addition to hypertension, cardiovascular abnormalities such as left ventricular hypertrophy, biventricular diastolic dysfunction, and impaired coronary flow
velocity reserve can develop. US is routinely used to rapidly measure common cardiovascular metrics like ejection fraction, stroke volume, and cardiac output. In future studies, it would be possible to measure TKV along with these cardiac metrics using the robotic US scanner evaluated herein.

One issue that is often problematic with in vivo imaging is respiratory artifacts. In this study animal respiration was not found to cause significant artifacts in US kidney scans. In Figure 1-B1 horizontal lines are evident on the skin of the mouse where breaths occurred, but are miniscule on the kidneys. This is because firstly, animals laid supine so the kidney displacement from lung expansion was negligible. Secondly, a high and low frequency scan were overlaid for segmentation, so the respiratory lines became less sharp. Based on these results, it does not appear that active respiration gating is necessary for accurate TKV measurement, which allows for a less complicated workflow and higher animal throughput.

There were several limitations to this study. First, the MRI and US imaging was not performed at the same facility requiring mice to be shipped and acclimated, which incurred a time delay between imaging modalities. While this delay was relatively short (20 days), it is conceivable that the kidney size may have changed due to natural biological progression and reduced the agreement between MRI and US. Second, the MRI and US datasets were evaluated by different readers using different segmentation software packages, with no explicit training between the two. It is expected that with improved training (e.g. developing a segmentation rubric), the correlation between MRI and US could increase beyond that reported in this study. Third, ex vivo kidney volume measurements were observed to have a time dependence with time of euthanasia (Supplemental Figure S5). In this study, all mice in Cohort 2 were euthanized via isoflurane overdose followed by cervical dislocation, with kidney extraction times varying depending on
the speed of the technician performing the necropsies. The first kidneys extracted had \textit{ex vivo} volumes that measured smaller than \textit{in vivo} volume, while the last kidneys extracted (~30 min post euthanasia) had \textit{ex vivo} volumes larger than the \textit{in vivo} volume. This indicates that post-mortem changes may have had an impact on the correlation between \textit{in vivo} and \textit{ex vivo} measurements and highlights the importance of necropsy standardization. Fourth, this study only evaluated a small number of mice (N=30 kidneys between the two cohorts), so larger studies with a broader range of kidney sizes and disease pathologies (e.g. kidney fibrosis models) should be conducted in the future. Finally, serum and urinary biomarkers with PKD progression prediction similar to TKV, such as β2MG and MCP-1 are not measured in this study but would provide more detailed disease stratification in future studies.

\textbf{Disclosures}

Several authors are either employed by, have a significant financial interest in, or are co-inventors on patents licensed by SonoVol, Inc. (NJB, RCG, JDR, PAD, TJC). Authors report NIH grant with SonoVol under consideration: R43 DK126607- (SBIR/NIH) (MPI; SonoVol PI - T Czernuszewicz/ Mayo; Mayo MPI - T Kline & MF Romero), A new robotic AI imaging platform for improved kidney disease research and drug discovery. This proposal builds and extends the work presented in this manuscript.

\textbf{Funding}

None
Acknowledgements

We would also like to acknowledge B. Velasco and the University of North Carolina Animal Core for their assistance with animal handling during the in vivo experiments.

Author Contributions

N Beaumont: Formal analysis; Visualization; Writing - original draft; Writing - review and editing

H Holmes: Data curation; Formal analysis; Writing - review and editing

A Gregory: Data curation; Formal analysis; Software

M Edwards: Data curation; Formal analysis

J Rojas: Formal analysis; Writing - review and editing

R Gessner: Conceptualization; Formal analysis; Funding acquisition; Resources; Supervision; Writing - review and editing

P Dayton: Resources; Supervision; Writing - review and editing

T Kline: Conceptualization; Resources; Supervision; Writing - review and editing

M Romero: Conceptualization; Funding acquisition; Project administration; Resources; Supervision; Writing - review and editing

T Czernuszewicz: Conceptualization; Data curation; Formal analysis; Funding acquisition; Methodology; Resources; Supervision; Writing - review and editing
References

1. Levy M, Feingold J: Estimating prevalence in single-gene kidney diseases progressing to renal failure. *Kidney Int.* [Internet] 58: 925–943, 2000 Available from: https://linkinghub.elsevier.com/retrieve/pii/S0085253815471844

2. Willey CJ, Blais JD, Hall AK, Krasa HB, Makin AJ, Czerwiec FS: Prevalence of autosomal dominant polycystic kidney disease in the European Union. *Nephrol. Dial. Transplant.* [Internet] gfw240, 2016 Available from: http://ndt.oxfordjournals.org/lookup/doi/10.1093/ndt/gfw240

3. Igarashi P, Somlo S: Polycystic Kidney Disease. *J. Am. Soc. Nephrol.* [Internet] 18: 1371–1373, 2007 Available from: http://www.jasn.org/lookup/doi/10.1681/ASN.2007030299

4. Sans-Atxer L, Joly D: Tolvaptan in the treatment of autosomal dominant polycystic kidney disease: patient selection and special considerations. *Int. J. Nephrol. Renovasc. Dis.* [Internet] Volume 11: 41–51, 2018 Available from: https://www.dovepress.com/tolvaptan--in-the-treatment-of-autosomal-dominant-polycystic-kidney-di-peer-reviewed-article-IJNRD

5. Nagao S, Kugita M, Yoshihara D, Yamaguchi T: Animal models for human polycystic kidney disease. *Exp. Anim.* 61: 477–88, 2012

6. Happé H, Peters DJM: Translational research in ADPKD: lessons from animal models. *Nat. Rev. Nephrol.* [Internet] 10: 587–601, 2014 Available from: http://www.nature.com/articles/nrneph.2014.137

7. Perrone RD, Mouksassi M-S, Romero K, Czerwiec FS, Chapman AB, Gitomer BY, Torres VE, Miskulin DC, Broadbent S, Marier JF: 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. Reports* [Internet] 2: 442–450, 2017 Available from: https://linkinghub.elsevier.com/retrieve/pii/S2468024917300037

8. Grantham JJ, Torres VE: The importance of total kidney volume in evaluating progression of polycystic kidney disease. *Nat. Rev. Nephrol.* 12: 667–677, 2016

9. Qualification of Biomarker Total Kidney Volume in Studies for Treatment of Autosomal Dominant Polycystic Kidney Disease Guidance for Industry [Internet]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/qualification-biomarker-total-kidney-volume-studies-treatment-autosomal-dominant-polycystic-kidney

10. Bidar AW, Ploj K, Lelliott C, Nelander K, Winzell MS, Böttcher G, Oscarsson J, Storlien L, Hockings PD: In vivo imaging of lipid storage and regression in diet-induced obesity during nutrition manipulation. *Am. J. Physiol. Metab.* [Internet] 303: E1287–E1295, 2012 Available from: https://www.physiology.org/doi/10.1152/ajpendo.00274.2012

11. Tangri N, Hougen I, Alam A, Perrone R, McFarlane P, Pei Y: Total kidney volume as a biomarker of disease progression in autosomal dominant polycystic kidney disease. *Can. J. Kidney Heal. Dis.* 4: 2017
12. Edwards ME, Blais JD, Czerwiec FS, Erickson BJ, Torres VE, Kline TL: Standardizing total kidney volume measurements for clinical trials of autosomal dominant polycystic kidney disease. *Clin. Kidney J.* [Internet] 12: 71–77, 2019 Available from: https://academic.oup.com/ckj/article/12/1/71/5086494

13. Kline TL, Edwards ME, Garg I, Irazabal M V, Korfiatis P, Harris PC, King BF, Torres VE, Venkatesh SK, Erickson BJ: Quantitative MRI of kidneys in renal disease. *Abdom. Radiol. (New York)* 43: 629–638, 2018

14. Wallace DP, Hou Y-P, Huang ZL, Nivens E, Savinkova L, Yamaguchi T, Bilgen M: Tracking kidney volume in mice with polycystic kidney disease by magnetic resonance imaging. *Kidney Int.* 73: 778–81, 2008

15. Czernuszewicz TJ, Papadopoulou V, Rojas JD, Rajamahendiran RM, Perdomo J, Butler J, Harlacher M, O’Connell G, Zukić D, Aylward SR, Dayton PA, Gessner RC: A new preclinical ultrasound platform for widefield 3D imaging of rodents. *Rev. Sci. Instrum.* 89: 075107, 2018

16. Xu R, Franchi F, Miller B, Crane JA, Peterson KM, Psaltis PJ, Harris PC, Lerman LO, Rodriguez-Porcel M: Polycystic Kidneys Have Decreased Vascular Density: A Micro-CT Study. *Microcirculation* 20: 183–189, 2013

17. Ogunlade O, Connell JJ, Huang JL, Zhang E, Lythgoe MF, Long DA, Beard P: In vivo three-dimensional photoacoustic imaging of the renal vasculature in preclinical rodent models. *Am. J. Physiol. - Ren. Physiol.* 314: F1145–F1153, 2018

18. Gessner RC, Frederick CB, Foster FS, Dayton PA: Acoustic angiography: A new imaging modality for assessing microvasculature architecture. *Int. J. Biomed. Imaging* 2013: 2013

19. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller J V., Pieper S, Kikinis R: 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* 30: 1323–1341, 2012

20. Kline TL, Edwards ME, Korfiatis P, Akkus Z, Torres VE, Erickson BJ: Semiautomated segmentation of polycystic kidneys in T2-weighted MR images. *Am. J. Roentgenol.* 207: 605–613, 2016

21. McGraw KO, Wong SP: Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* 1: 30–46, 1996

22. Chao K, Liao K, Khan M, Shi C, Li J, Goldberg ID, Narayan P: An Improved Method for Estimating Renal Dimensions; Implications for Management of Kidney Disease. *Appl. Sci.* 9: 3198, 2019

23. Chapman AB, Stepniakowski K, Rahbari-Oskoui F: Hypertension in Autosomal Dominant Polycystic Kidney Disease. *Adv. Chronic Kidney Dis.* [Internet] 17: 153–163, 2010 Available from: https://linkinghub.elsevier.com/retrieve/pii/S1548559510000029

24. Bergmann C: ARPKD and early manifestations of ADPKD: the original polycystic kidney disease and phenocopies. *Pediatr. Nephrol.* [Internet] 30: 15–30, 2015 Available from: http://link.springer.com/10.1007/s00467-013-2706-2
25. Chapman AB, Guay-Woodford LM, Grantham JJ, Torres VE, Bae KT, Baumgarten DA, Kenney PJ, King BF, Glockner JF, Wetzel LH, Brummer ME, Charles O’Neill W, Robbin ML, Bennett WM, Klahr S, Hirschman GH, Kimmel PL, Thompson PA, Philip Miller J: Renal structure in early autosomal-dominant polycystic kidney disease (ADPKD): The Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP) cohort1. Kidney Int. [Internet] 64: 1035–1045, 2003 Available from: https://linkinghub.elsevier.com/retrieve/pii/S0085253815494256

26. Rahman E, Niaz FA, Al-suwaida A, Nahrir S, Bashir M, Rahman H, Hammad D: Polycystic Kidney Disease: A Single Center Study. Nephrology 20: 806–810, 2009

27. Ecder T, Schrier RW: Cardiovascular abnormalities in autosomal-dominant polycystic kidney disease. Nat. Rev. Nephrol. [Internet] 5: 221–228, 2009 Available from: http://www.nature.com/articles/nrneph.2009.13

28. Messchendorp AL, Meijer E, Visser FW, Engels GE, Kappert P, Losekoot M, Peters DJM, Gansevoort RT: Rapid Progression of Autosomal Dominant Polycystic Kidney Disease: Urinary Biomarkers as Predictors. Am. J. Nephrol. [Internet] 50: 375–385, 2019 Available from: https://www.karger.com/Article/FullText/502999

29. Rojas JD, Papadopoulou V, Czernuszewicz TJ, Rajamahendiran RM, Chytil A, Chiang YC, Chong DC, Bautch VL, Rathmell WK, Aylward S, Gessner RC, Dayton PA: Ultrasound Measurement of Vascular Density to Evaluate Response to Anti-Angiogenic Therapy in Renal Cell Carcinoma. IEEE Trans. Biomed. Eng. 66: 873–880, 2019
Figure 1: An overview of the acquisition workflow for the robotic ultrasound system. Multiple parallel sweeps are acquired, usually three, resulting in thousands of individual 2D images from across the animal’s body (A). These images are then stitched together to produce a single widefield 3D image volume, which can be viewed in different orientations (B, C). The 3D data can be seen in frontal (B1), transverse (B2), and sagittal (B3) planes, respectively. Each kidney can then be segmented in 3D to assess TKV (C). Scale bar indicates 1 cm.
Figure 2: Robotic US and MRI images of the same kidney in the same slices show cysts (yellow arrows) in the same regions. The inferior vena cava is also visible (red arrow).
Figure 3: Regression and Bland-Altman plots comparing robotic ultrasound-based *in vivo* measurements of kidney volume to *in vivo* MRI. $r$, Pearson correlation coefficient; $r^2$, Pearson’s coefficient squared; $R^2$, coefficient of determination; $\rho$, Spearman correlation coefficient; $LOA$, limits of agreement; $CV$, coefficient of variation.
Figure 4: A) An image showing how Vernier calipers were used to measure kidney length (dashed line) and width (solid line). B) US calipers drawn in software over a coronal *in vivo* US image. C) US calipers drawn over an US image of *ex vivo* kidneys. Lines in A, B and C correspond to the same axes. D-F) Bland-Altman plots comparing *in vivo* US calipers, *ex vivo* US calipers, and *ex vivo* Vernier calipers.
Figure 5: Inter-reader variability for Total Kidney Volume (TKV) measurements for Cohort 1.

Readers are indicated by the marker, with mean and standard deviations plotted.
Figure 6

A. Maximum intensity projection (MIP) through 3D AA volume

B. Aggregate Histogram of AA Images

C. Correlation: Kidney Vascularity and Size

D. Linear Regression $R^2$ vs. Threshold

$R^2 = 0.82$
Figure 6: Correlation between kidney vascularity and kidney size. A) Vessel morphology as visualized via maximum intensity projections of AA volumes. A single 2D slice of the volume with an exemplary cyst (white arrow) is shown. After thresholding at the optimal threshold of 46 counts, the voxels within the cyst, and elsewhere, have been set to zero and not included in the % positivity metric for vascularity. B) Mean histogram of all AA volumes of Cohort 1. Standard deviations of the histogram bins are shown with error bars. C) The percent of bright voxels in a kidney during microbubble perfusion (% positivity) negatively correlates with kidney volume. In this plot the kidneys were thresholded at 46 counts. D) The coefficient of determination of the linear regression in C) depends on the thresholding level. Thresholding at 46 counts gives the highest R2 value. This threshold excludes cysts but includes most vasculature (seen on the right in panel A).