Table S1. Clinical and laboratory findings of five control subjects each with 1-3 renal cysts on ultrasound

Subject	Sex	Age at US and MRI	Total # of cysts on US	Max. cyst size (mm)	Date of BP & kidney tests	Current age (yrs)	sCr (mg/dL)	BP	Urinalysis
1	M	29.7	1	5	2-Jun-14	32.8	0.90	123/75	Normal
2	F	28.5	3	31	5-Jun-14	32	0.73	124/83	Normal
3	M	39.4	1	2.6	16-Jun-14	43	1.00	122/82	Normal
4	M	31.2	2	3.9	4-Jun-14	34.3	0.90	112/80	Normal
5	M	24.8	1	3.9	30-May-14	27	0.90	118/76	Normal

*Normal urinalysis: negative dipstick testing for blood, protein, and glucose and normal microscopy.

Inter-rater Reliability of MRI Cyst Counts in Subjects at Risk for Autosomal Dominant Polycystic Kidney Disease (ADPKD)

J. Conklin, MD, S.T. Schindera, MD, D. Mercer, MD, Y. Pei, MD, M.A. Haider, MD

From the Department of Medical Imaging (J.C., S.T.S., D.M., M.A.H.) and the

Division of Nephrology (Y.P.), University Health Network, Toronto, Ontario, Canada.

Methods:

Subjects were participants in an ongoing study to determine optimal MRI criteria for diagnosis of ADPKD, for which informed consent and institutional review ethics board approval had been obtained. From this cohort, all subjects aged 16 to 29 were selected for the retrospective image review and analysis procedure described below. Subjects were members of known ADPKD families (i.e., 50% risk of inheriting either PKD1 or PKD2 at birth) or age matched controls with no history of systemic or genitourinary disease. A total of 87 subjects (57 at risk individuals and 30 controls) were selected to provide an approximately equal number of mutation positive and mutation negative individuals.

Imaging was performed on one of two 1.5T MRI systems (Signa HDxt, GE Medical Systems, or Magnetom Avanto, Siemens) using an 8-channel phased array body coil. The following image series were acquired in both axial and coronal planes, with sufficient coverage to include the full anteroposterior and craniocaudal extent of the both kidneys: 1) T2-weighted single-shot fast spin-echo (SS-FSE) with fat suppression (TE = 180 – 300ms, TR = 3500 ms, flip angle (FA) = 90°, FOV = 300×300 mm, matrix = 256×256, slice thickness = 3mm, slice gap = 0, acceleration factor = 2); 2) T2-weighted SS-FSE without fat suppression (TE = 90ms, TR = 2500 ms, FA = 90°, FOV = 300×300 mm, matrix = 256×256, slice thickness = 3mm, slice gap = 0, acceleration factor = 2); and, 3) 3D T1-weighted fast spoiled gradient-echo with fat suppression (TE = 2.3 ms, TR = 5.2 ms, FA = 12°, FOV = 300×300 mm, matrix = 256×256, slice thickness = 3mm, slice gap = 0, acceleration factor = 2).

All studies were reviewed by 3 independent readers. One reader was a fellowship trained radiologist with 15 years of experience in abdominal imaging and the other two readers were board certified radiologists undergoing fellowship training in abdominal imaging. Studies were reviewed in randomized sequence on a standard PACS workstation (eFilm Workstation, Merge Healthcare, or ClearCanvas Workstation, ClearCanvas Inc). At a minimum, the axial fat

suppressed T2-weighted images were reviewed, with correlation to additional series at the reader's discretion. Cysts were counted and categorized as small (≤10mm) or large (>10mm), up to a maximum of 10 cysts per kidney. Parapelvic cysts were not counted due to the difficulty in differentiating between these cysts and the collecting system on T2-weighted images. The maximum cyst diameter within each kidney was also recorded. Cyst counts were then binned into the following categories: 0 to 1, 2 to 4, 5 to 10, or greater than 10 cysts per kidney.

For binned categorical data, Cohen's unweighted kappa was calculated for each pair of readers, as was Fleiss multiple kappa for overall agreement between all readers.[2] For all kidneys with 0 to 10 cysts, intra-class correlation coefficients (ICCs) were calculated for the exact number of small, large, and total cysts. The effects of gender, age, and laterality (right versus left kidney) on total number of cysts (using the binned categories defined above) was assessed using univariate ordered logistic regression tests, with the categorical cyst count as the dependent variable.

Finally, two subgroups of kidneys were identified to specifically assess inter-rater reliability in cases where ultrasound findings may be equivocal, specifically: (1) all kidneys with at least 1 cyst and a maximum cyst dimension ≤ 10 mm, and (2) all kidneys at least 1 cyst and a maximum cyst dimension ≤ 5mm. The unweighted kappa for the three rater pairs and the ICC for the total number of cysts were calculated for these two subsets of kidneys.

Results:

Of the 174 kidneys in 87 subjects, 78 kidneys in 49 subjects had at least one renal cyst identified by at least one reader. The distribution of kidneys by cyst count and maximum cyst dimension is shown in Table 1. There was no significant effect of gender or age on the total number of cysts, nor was there a significant difference between cyst counts for the left and right kidneys.

Kappa values for categorical cyst counts were 0.96, 0.97, and 0.96 for the 3 pairs of readers, representing almost perfect agreement by the standard interpretation of Landis and Koch.[1] Fleiss multiple kappa for agreement between all 3 readers was 0.97 (95% confidence interval 0.95 to

0.99). The ICCs for quantitative cyst counts of the 3 readers were 0.95 for small cysts, 0.96 for large cysts, and 0.96 for total cyst number.

Next we considered only those kidneys with a maximum cyst diameter ≤ 10 mm, including all such kidneys with at least one cyst identified by any rater. For these 26 kidneys, kappa values for binned cyst counts were 0.92, 0.86, and 0.78, with an overall multiple kappa of 0.85 (95% confidence interval 0.81 to 0.89). The ICC for the exact cyst counts of the 3 readers was 0.95.

Finally we considered only those kidneys with a maximum cyst diameter ≤ 5 mm, including all such kidneys with at least one cyst identified by any rater. For these 18 kidneys, kappa values were 0.83, 0.85, and 0.83, with an overall multiple kappa of 0.79 (95% confidence interval 0.74 to 0.84). The ICC for the exact cyst counts of the 3 readers was 0.76.

Conclusions:

Inter-observer agreement of renal cyst counts by MRI is excellent across a range of cyst size to less than 5 mm in diameter.

References:

1. Landis JR, Koch CG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159-174.

2. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York, NY: Wiley, 1981; 212-236.

Table 1. Distribution of cyst count and maximum cyst diameter from 174 Kidneys

Cysts/Kidney	Maximum Cyst Dimension (mm)			
	≤ 5mm	5–10 mm	> 10mm	Total
0 cysts	-	-	-	102
1 cyst	9	3	3	15
2–4 cysts	0	2	2	4
5–10 cysts	0	2	1	3
> 10 cysts	2	2	46	50
Total	11	9	52	174

Note: Data are the number of kidneys with each total cyst count and maximum cyst dimension, based on median cyst count of the 3 readers.